

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
WO 02/20754 A2

(51) International Patent Classification<sup>7</sup>: C12N 15/00

(21) International Application Number: PCT/US01/27127

(22) International Filing Date: 29 August 2001 (29.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/229,751	5 September 2000 (05.09.2000)	US
60/229,749	5 September 2000 (05.09.2000)	US
60/229,750	5 September 2000 (05.09.2000)	US
60/229,747	5 September 2000 (05.09.2000)	US
60/229,748	5 September 2000 (05.09.2000)	US
60/230,583	5 September 2000 (05.09.2000)	US
60/230,517	6 September 2000 (06.09.2000)	US
60/230,599	6 September 2000 (06.09.2000)	US
60/230,514	6 September 2000 (06.09.2000)	US
60/230,598	6 September 2000 (06.09.2000)	US
60/230,988	6 September 2000 (06.09.2000)	US
60/230,518	6 September 2000 (06.09.2000)	US
60/230,515	6 September 2000 (06.09.2000)	US
60/230,597	6 September 2000 (06.09.2000)	US
60/230,505	6 September 2000 (06.09.2000)	US
60/230,519	6 September 2000 (06.09.2000)	US
60/230,595	6 September 2000 (06.09.2000)	US
60/230,865	6 September 2000 (06.09.2000)	US
60/230,610	6 September 2000 (06.09.2000)	US
60/231,167	7 September 2000 (07.09.2000)	US
60/231,163	7 September 2000 (07.09.2000)	US
60/230,951	7 September 2000 (07.09.2000)	US

(71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): STUART, Jackson [US/US]; 4913 Deep Creek Road, Fremont, CA 94555 (US). LINCOLN, Stephen, E. [US/US]; 725 Sapphire Street, Redwood City, CA 94061 (US). ALTUS, Christina, M. [US/US]; 625 Virginia Avenue, Campbell, CA 95008 (US). DUFOUR, Gerard, E. [US/US];

5327 Greenridge Road, Castro Valley, CA 94552 (US). CHALUP, Michael, S. [US/US]; 183 Acalanes Drive, Apt. 6, Sunnyvale, CA 94086 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monrow Drive, #17, Mountain View, CA 94040 (US). JONES, Anissa, Lee [US/US]; 445 South 15th Street, San Jose, CA 95112 (US). YU, Jimmy, Y. [US/US]; 37330 Portico Terrace, Fremont, CA 94536 (US). WRIGHT, Rachel, J. [NZ/US]; 333 Anna Avenue, Mountain View, CA 94043 (US). GIETZEN, Darryl [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). LIU, Tommy, F. [US/US]; 201 Otilia Street, Daly city, Ca 94014 (US). YAP, Pierre, E. [US/US]; 201 Happy Hollow Court, Lafayette, CA 94549 (US). DAHL, Christopher, R. [US/US]; 41277 Roberts Avenue, #6, Fremont, CA 94538 (US). MOMIYAMA, Monika, G. [US/US]; 25689 Deer Trail Place, Hayward, CA 94541 (US). BRADLEY, Diana, L. [US/US]; 3260 Crystal Heights Drive, Soquel, CA 95073 (US). ROHATGI, Sameer, D. [US/US]; 5 Rico Way #103, San Francisco, CA 94123 (US). HARRIS, Bernard [US/US]; 1014 Lupine Drive, Sunnyvale, CA 94086 (US). ROSEBERRY, Ann, M. [US/US]; 725 Sapphire Street, Redwood City, CA 94061 (US). GERSTIN, Edward, H., Jr. [US/US]; 1408 38th Avenue, San Francisco, CA 94122 (US). PERALTA, Careyna, H. [US/US]; 4585 Lakeshore Drive, Santa Clara, CA 95054 (US). DAVID, Marie, H. [US/US]; 131 Mirada Drive, Daly City, CA 94015 (US). PANZER, Scott, R. [US/US]; 965 East El Camino, #621, Sunnyvale, CA 94087 (US). FLORES, Vincent [US/US]; 35000 Begonia Street, Union City, CA 94587 (US). DAFFO, Abel [US/US]; 1750 Stokes Street, #70, San Jose, CA 95126 (US). MARWAHA, Rakesh [US/US]; 16272 Saratoga Street, #4, San Leandro, CA 94578 (US). CHEN, Alice, J. [US/US]; 4405 Norwalk Drive, #22, San Jose, CA 95129 (US). CHANG, Simon, C. [US/US]; 1901 Rock Street, #103, Mountain View, CA 94043 (US). AU, Alan, P. [US/US]; 565 Ortega Avenue, #3, Mountain View, CA 94040 (US). INMAN, Rebekah, R. [US/US]; 210 Calderon Avenue, #23, Mountain View, CA 94041 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,

[Continued on next page]

(54) Title: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS

(57) Abstract: The present invention provides purified human polynucleotides for diagnostics and therapeutics (dithp). Also encompassed are the polypeptides (DITHP) encoded by dithp. The invention also provides for the use of dithp, or complements, oligonucleotides, or fragments thereof in diagnostic assays. The invention further provides for vectors and host cells containing dithp for the expression of DITHP. The invention additionally provides for the use of isolated and purified DITHP to induce antibodies and to screen libraries of compounds and the use of anti-DITHP antibodies in diagnostic assays. Also provided are microarrays containing dithp and methods of use.

WO 02/20754 A2



CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

**(84) Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



## MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS

### TECHNICAL FIELD

5           The present invention relates to human molecules and to the use of these sequences in the diagnosis, study, prevention, and treatment of diseases associated with, as well as effects of exogenous compounds on, the expression of human molecules.

### BACKGROUND OF THE INVENTION

10           The human genome is comprised of thousands of genes, many encoding gene products that function in the maintenance and growth of the various cells and tissues in the body. Aberrant expression or mutations in these genes and their products is the cause of, or is associated with, a variety of human diseases such as cancer and other cell proliferative disorders, autoimmune/inflammatory disorders, infections, developmental disorders, endocrine disorders,  
15           metabolic disorders, neurological disorders, gastrointestinal disorders, transport disorders, and connective tissue disorders. The identification of these genes and their products is the basis of an ever-expanding effort to find markers for early detection of diseases, and targets for their prevention and treatment. Therefore, these genes and their products are useful as diagnostics and therapeutics. These genes may encode, for example, enzyme molecules, molecules associated with growth and  
20           development, biochemical pathway molecules, extracellular information transmission molecules, receptor molecules, intracellular signaling molecules, membrane transport molecules, protein modification and maintenance molecules, nucleic acid synthesis and modification molecules, adhesion molecules, antigen recognition molecules, secreted and extracellular matrix molecules, cytoskeletal molecules, ribosomal molecules, electron transfer associated molecules, transcription  
25           factor molecules, chromatin molecules, cell membrane molecules, and organelle associated molecules.

          For example, cancer represents a type of cell proliferative disorder that affects nearly every tissue in the body. A wide variety of molecules, either aberrantly expressed or mutated, can be the cause of, or involved with, various cancers because tissue growth involves complex and ordered  
30           patterns of cell proliferation, cell differentiation, and apoptosis. Cell proliferation must be regulated to maintain both the number of cells and their spatial organization. This regulation depends upon the appropriate expression of proteins which control cell cycle progression in response to extracellular signals such as growth factors and other mitogens, and intracellular cues such as DNA damage or nutrient starvation. Molecules which directly or indirectly modulate cell cycle progression fall into  
35           several categories, including growth factors and their receptors, second messenger and signal transduction proteins, oncogene products, tumor-suppressor proteins, and mitosis-promoting factors. Aberrant expression or mutations in any of these gene products can result in cell proliferative

disorders such as cancer. Oncogenes are genes generally derived from normal genes that, through abnormal expression or mutation, can effect the transformation of a normal cell to a malignant one (oncogenesis). Oncoproteins, encoded by oncogenes, can affect cell proliferation in a variety of ways and include growth factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and cell-cycle control proteins. In contrast, tumor-suppressor genes are involved in inhibiting cell proliferation. Mutations which cause reduced function or loss of function in tumor-suppressor genes result in aberrant cell proliferation and cancer. Although many different genes and their products have been found to be associated with cell proliferative disorders such as cancer, many more may exist that are yet to be discovered.

DNA-based arrays can provide a simple way to explore the expression of a single polymorphic gene or a large number of genes. When the expression of a single gene is explored, DNA-based arrays are employed to detect the expression of specific gene variants. For example, a p53 tumor suppressor gene array is used to determine whether individuals are carrying mutations that predispose them to cancer. A cytochrome p450 gene array is useful to determine whether individuals have one of a number of specific mutations that could result in increased drug metabolism, drug resistance or drug toxicity.

DNA-based array technology is especially relevant for the rapid screening of expression of a large number of genes. There is a growing awareness that gene expression is affected in a global fashion. A genetic predisposition, disease or therapeutic treatment may affect, directly or indirectly, the expression of a large number of genes. In some cases the interactions may be expected, such as when the genes are part of the same signaling pathway. In other cases, such as when the genes participate in separate signaling pathways, the interactions may be totally unexpected. Therefore, DNA-based arrays can be used to investigate how genetic predisposition, disease, or therapeutic treatment affects the expression of a large number of genes.

### **Enzyme Molecules**

The cellular processes of biogenesis and biodegradation involve a number of key enzyme classes including oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases. These enzyme classes are each comprised of numerous substrate-specific enzymes having precise and well regulated functions. These enzymes function by facilitating metabolic processes such as glycolysis, the tricarboxylic cycle, and fatty acid metabolism; synthesis or degradation of amino acids, steroids, phospholipids, alcohols, etc.; regulation of cell signalling, proliferation, inflammation, apoptosis, etc., and through catalyzing critical steps in DNA replication and repair, and the process of translation.

#### **Oxidoreductases**

Many pathways of biogenesis and biodegradation require oxidoreductase (dehydrogenase or reductase) activity, coupled to the reduction or oxidation of a donor or acceptor cofactor. Potential

cofactors include cytochromes, oxygen, disulfide, iron-sulfur proteins, flavin adenine dinucleotide (FAD), and the nicotinamide adenine dinucleotides NAD and NADP (Newsholme, E.A. and A.R. Leech (1983) Biochemistry for the Medical Sciences, John Wiley and Sons, Chichester, U.K., pp. 779-793). Reductase activity catalyzes the transfer of electrons between substrate(s) and cofactor(s)

5 with concurrent oxidation of the cofactor. The reverse dehydrogenase reaction catalyzes the reduction of a cofactor and consequent oxidation of the substrate. Oxidoreductase enzymes are a broad superfamily of proteins that catalyze numerous reactions in all cells of organisms ranging from bacteria to plants to humans. These reactions include metabolism of sugar, certain detoxification reactions in the liver, and the synthesis or degradation of fatty acids, amino acids, glucocorticoids, estrogens, androgens, and prostaglandins. Different family members are named according to the direction in which their reactions are typically catalyzed; thus they may be referred to as oxidoreductases, oxidases, reductases, or dehydrogenases. In addition, family members often have distinct cellular localizations, including the cytosol, the plasma membrane, mitochondrial inner or outer membrane, and peroxisomes.

15 Short-chain alcohol dehydrogenases (SCADs) are a family of dehydrogenases that only share 15% to 30% sequence identity, with similarity predominantly in the coenzyme binding domain and the substrate binding domain. In addition to the well-known role in detoxification of ethanol, SCADs are also involved in synthesis and degradation of fatty acids, steroids, and some prostaglandins, and are therefore implicated in a variety of disorders such as lipid storage disease, myopathy, SCAD deficiency, and certain genetic disorders. For example, retinol dehydrogenase is a SCAD-family member (Simon, A. et al. (1995) *J. Biol. Chem.* 270:1107-1112) that converts retinol to retinal, the precursor of retinoic acid. Retinoic acid, a regulator of differentiation and apoptosis, has been shown to down-regulate genes involved in cell proliferation and inflammation (Chai, X. et al. (1995) *J. Biol. Chem.* 270:3900-3904). In addition, retinol dehydrogenase has been linked to hereditary eye diseases 20 such as autosomal recessive childhood-onset severe retinal dystrophy (Simon, A. et al. (1996) *Genomics* 36:424-430).

Propagation of nerve impulses, modulation of cell proliferation and differentiation, induction of the immune response, and tissue homeostasis involve neurotransmitter metabolism (Weiss, B. (1991) *Neurotoxicology* 12:379-386; Collins, S.M. et al. (1992) *Ann. N.Y. Acad. Sci.* 664:415-424; 30 Brown, J.K. and H. Imam (1991) *J. Inherit. Metab. Dis.* 14:436-458). Many pathways of neurotransmitter metabolism require oxidoreductase activity, coupled to reduction or oxidation of a cofactor, such as NAD<sup>+</sup>/NADH (Newsholme, E.A. and A.R. Leech (1983) Biochemistry for the Medical Sciences, John Wiley and Sons, Chichester, U.K. pp. 779-793). Degradation of catecholamines (epinephrine or norepinephrine) requires alcohol dehydrogenase (in the brain) or aldehyde dehydrogenase (in peripheral tissue). NAD<sup>+</sup>-dependent aldehyde dehydrogenase oxidizes 35 5-hydroxyindole-3-acetate (the product of 5-hydroxytryptamine (serotonin) metabolism) in the brain,

blood platelets, liver and pulmonary endothelium (Newsholme, supra, p. 786). Other neurotransmitter degradation pathways that utilize NAD<sup>+</sup>/NADH-dependent oxidoreductase activity include those of L-DOPA (precursor of dopamine, a neuronal excitatory compound), glycine (an inhibitory neurotransmitter in the brain and spinal cord), histamine (liberated from mast cells during the inflammatory response), and taurine (an inhibitory neurotransmitter of the brain stem, spinal cord and retina) (Newsholme, supra, pp. 790, 792). Epigenetic or genetic defects in neurotransmitter metabolic pathways can result in a spectrum of disease states in different tissues including Parkinson disease and inherited myoclonus (McCance, K.L. and S.E. Huether (1994) Pathophysiology, Mosby-Year Book, Inc., St. Louis MO, pp. 402-404; Gundlach, A.L. (1990) FASEB J. 4:2761-2766).

Tetrahydrofolate is a derivatized glutamate molecule that acts as a carrier, providing activated one-carbon units to a wide variety of biosynthetic reactions, including synthesis of purines, pyrimidines, and the amino acid methionine. Tetrahydrofolate is generated by the activity of a holoenzyme complex called tetrahydrofolate synthase, which includes three enzyme activities: tetrahydrofolate dehydrogenase, tetrahydrofolate cyclohydrolase, and tetrahydrofolate synthetase. Thus, tetrahydrofolate dehydrogenase plays an important role in generating building blocks for nucleic and amino acids, crucial to proliferating cells.

3-Hydroxyacyl-CoA dehydrogenase (3HACD) is involved in fatty acid metabolism. It catalyzes the reduction of 3-hydroxyacyl-CoA to 3-oxoacyl-CoA, with concomitant oxidation of NAD to NADH, in the mitochondria and peroxisomes of eukaryotic cells. In peroxisomes, 3HACD and enoyl-CoA hydratase form an enzyme complex called bifunctional enzyme, defects in which are associated with peroxisomal bifunctional enzyme deficiency. This interruption in fatty acid metabolism produces accumulation of very-long chain fatty acids, disrupting development of the brain, bone, and adrenal glands. Infants born with this deficiency typically die within 6 months (Watkins, P. et al. (1989) J. Clin. Invest. 83:771-777; Online Mendelian Inheritance in Man (OMIM), #261515). The neurodegeneration that is characteristic of Alzheimer's disease involves development of extracellular plaques in certain brain regions. A major protein component of these plaques is the peptide amyloid- $\beta$  (A $\beta$ ), which is one of several cleavage products of amyloid precursor protein (APP). 3HACD has been shown to bind the A $\beta$  peptide, and is overexpressed in neurons affected in Alzheimer's disease. In addition, an antibody against 3HACD can block the toxic effects of A $\beta$  in a cell culture model of Alzheimer's disease (Yan, S. et al. (1997) Nature 389:689-695; OMIM, #602057).

Steroids, such as estrogen, testosterone, corticosterone, and others, are generated from a common precursor, cholesterol, and are interconverted into one another. A wide variety of enzymes act upon cholesterol, including a number of dehydrogenases. Steroid dehydrogenases, such as the hydroxysteroid dehydrogenases, are involved in hypertension, fertility, and cancer (Duax, W.L. and D. Ghosh (1997) Steroids 62:95-100). One such dehydrogenase is 3-oxo-5- $\alpha$ -steroid dehydrogenase

(OASD), a microsomal membrane protein highly expressed in prostate and other androgen-responsive tissues. OASD catalyzes the conversion of testosterone into dihydrotestosterone, which is the most potent androgen. Dihydrotestosterone is essential for the formation of the male phenotype during embryogenesis, as well as for proper androgen-mediated growth of tissues such as the prostate and male genitalia. A defect in OASD that prevents the conversion of testosterone into dihydrotestosterone leads to a rare form of male pseudohermaphroditis, characterized by defective formation of the external genitalia (Andersson, S. et al. (1991) *Nature* 354:159-161; Labrie, F. et al. (1992) *Endocrinology* 131:1571-1573; OMIM #264600). Thus, OASD plays a central role in sexual differentiation and androgen physiology.

17 $\beta$ -hydroxysteroid dehydrogenase (17 $\beta$ HSD6) plays an important role in the regulation of the male reproductive hormone, dihydrotestosterone (DHTT). 17 $\beta$ HSD6 acts to reduce levels of DHTT by oxidizing a precursor of DHTT, 3 $\alpha$ -diol, to androsterone which is readily glucuronidated and removed from tissues. 17 $\beta$ HSD6 is active with both androgen and estrogen substrates when expressed in embryonic kidney 293 cells. At least five other isozymes of 17 $\beta$ HSD have been identified that catalyze oxidation and/or reduction reactions in various tissues with preferences for different steroid substrates (Biswas, M.G. and D.W. Russell (1997) *J. Biol. Chem.* 272:15959-15966). For example, 17 $\beta$ HSD1 preferentially reduces estradiol and is abundant in the ovary and placenta. 17 $\beta$ HSD2 catalyzes oxidation of androgens and is present in the endometrium and placenta. 17 $\beta$ HSD3 is exclusively a reductive enzyme in the testis (Geissler, W.M. et al. (1994) *Nat. Genet.* 7:34-39). An excess of androgens such as DHTT can contribute to certain disease states such as benign prostatic hyperplasia and prostate cancer.

Oxidoreductases are components of the fatty acid metabolism pathways in mitochondria and peroxisomes. The main beta-oxidation pathway degrades both saturated and unsaturated fatty acids, while the auxiliary pathway performs additional steps required for the degradation of unsaturated fatty acids. The auxiliary beta-oxidation enzyme 2,4-dienoyl-CoA reductase catalyzes the removal of even-numbered double bonds from unsaturated fatty acids prior to their entry into the main beta-oxidation pathway. The enzyme may also remove odd-numbered double bonds from unsaturated fatty acids (Koivuranta, K.T. et al. (1994) *Biochem. J.* 304:787-792; Smeland, T.E. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6673-6677). 2,4-dienoyl-CoA reductase is located in both mitochondria and peroxisomes. Inherited deficiencies in mitochondrial and peroxisomal beta-oxidation enzymes are associated with severe diseases, some of which manifest themselves soon after birth and lead to death within a few years. Defects in beta-oxidation are associated with Reye's syndrome, Zellweger syndrome, neonatal adrenoleukodystrophy, infantile Refsum's disease, acyl-CoA oxidase deficiency, and bifunctional protein deficiency (Suzuki, Y. et al. (1994) *Am. J. Hum. Genet.* 54:36-43; Hoefler, *supra*; Cotran, R.S. et al. (1994) *Robbins Pathologic Basis of Disease*, W.B. Saunders Co., Philadelphia PA, p.866). Peroxisomal beta-oxidation is impaired in cancerous tissue. Although

neoplastic human breast epithelial cells have the same number of peroxisomes as do normal cells, fatty acyl-CoA oxidase activity is lower than in control tissue (el Bouhtoury, F. et al. (1992) *J. Pathol.* 166:27-35). Human colon carcinomas have fewer peroxisomes than normal colon tissue and have lower fatty-acyl-CoA oxidase and bifunctional enzyme (including enoyl-CoA hydratase) activities than normal tissue (Cable, S. et al. (1992) *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* 62:221-226). Another important oxidoreductase is isocitrate dehydrogenase, which catalyzes the conversion of isocitrate to  $\alpha$ -ketoglutarate, a substrate of the citric acid cycle. Isocitrate dehydrogenase can be either NAD or NADP dependent, and is found in the cytosol, mitochondria, and peroxisomes. Activity of isocitrate dehydrogenase is regulated developmentally, and by hormones, neurotransmitters, and growth factors.

Hydroxypyruvate reductase (HPR), a peroxisomal 2-hydroxyacid dehydrogenase in the glycolate pathway, catalyzes the conversion of hydroxypyruvate to glycerate with the oxidation of both NADH and NADPH. The reverse dehydrogenase reaction reduces  $\text{NAD}^+$  and  $\text{NADP}^+$ . HPR recycles nucleotides and bases back into pathways leading to the synthesis of ATP and GTP. ATP and GTP are used to produce DNA and RNA and to control various aspects of signal transduction and energy metabolism. Inhibitors of purine nucleotide biosynthesis have long been employed as antiproliferative agents to treat cancer and viral diseases. HPR also regulates biochemical synthesis of serine and cellular serine levels available for protein synthesis.

The mitochondrial electron transport (or respiratory) chain is a series of oxidoreductase-type enzyme complexes in the mitochondrial membrane that is responsible for the transport of electrons from NADH through a series of redox centers within these complexes to oxygen, and the coupling of this oxidation to the synthesis of ATP (oxidative phosphorylation). ATP then provides the primary source of energy for driving a cell's many energy-requiring reactions. The key complexes in the respiratory chain are NADH:ubiquinone oxidoreductase (complex I), succinate:ubiquinone oxidoreductase (complex II), cytochrome  $c_1$ -b oxidoreductase (complex III), cytochrome c oxidase (complex IV), and ATP synthase (complex V) (Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, Inc., New York NY, pp. 677-678). All of these complexes are located on the inner matrix side of the mitochondrial membrane except complex II, which is on the cytosolic side. Complex II transports electrons generated in the citric acid cycle to the respiratory chain. The electrons generated by oxidation of succinate to fumarate in the citric acid cycle are transferred through electron carriers in complex II to membrane bound ubiquinone (Q). Transcriptional regulation of these nuclear-encoded genes appears to be the predominant means for controlling the biogenesis of respiratory enzymes. Defects and altered expression of enzymes in the respiratory chain are associated with a variety of disease conditions.

Other dehydrogenase activities using NAD as a cofactor are also important in mitochondrial function. 3-hydroxyisobutyrate dehydrogenase (3HBD), important in valine catabolism, catalyzes the

NAD-dependent oxidation of 3-hydroxyisobutyrate to methylmalonate semialdehyde within mitochondria. Elevated levels of 3-hydroxyisobutyrate have been reported in a number of disease states, including ketoacidosis, methylmalonic acidemia, and other disorders associated with deficiencies in methylmalonate semialdehyde dehydrogenase (Rougraff, P.M. et al. (1989) J. Biol. Chem. 264:5899-5903).

Another mitochondrial dehydrogenase important in amino acid metabolism is the enzyme isovaleryl-CoA-dehydrogenase (IVD). IVD is involved in leucine metabolism and catalyzes the oxidation of isovaleryl-CoA to 3-methylcrotonyl-CoA. Human IVD is a tetrameric flavoprotein that is encoded in the nucleus and synthesized in the cytosol as a 45 kDa precursor with a mitochondrial import signal sequence. A genetic deficiency, caused by a mutation in the gene encoding IVD, results in the condition known as isovaleric acidemia. This mutation results in inefficient mitochondrial import and processing of the IVD precursor (Vockley, J. et al. (1992) J. Biol. Chem. 267:2494-2501).

#### Transferases

Transferases are enzymes that catalyze the transfer of molecular groups. The reaction may involve an oxidation, reduction, or cleavage of covalent bonds, and is often specific to a substrate or to particular sites on a type of substrate. Transferases participate in reactions essential to such functions as synthesis and degradation of cell components, regulation of cell functions including cell signaling, cell proliferation, inflammation, apoptosis, secretion and excretion. Transferases are involved in key steps in disease processes involving these functions. Transferases are frequently classified according to the type of group transferred. For example, methyl transferases transfer one-carbon methyl groups, amino transferases transfer nitrogenous amino groups, and similarly denominated enzymes transfer aldehyde or ketone, acyl, glycosyl, alkyl or aryl, isoprenyl, saccharyl, phosphorous-containing, sulfur-containing, or selenium-containing groups, as well as small enzymatic groups such as Coenzyme A.

Acyl transferases include peroxisomal carnitine octanoyl transferase, which is involved in the fatty acid beta-oxidation pathway, and mitochondrial carnitine palmitoyl transferases, involved in fatty acid metabolism and transport. Choline O-acetyl transferase catalyzes the biosynthesis of the neurotransmitter acetylcholine.

Amino transferases play key roles in protein synthesis and degradation, and they contribute to other processes as well. For example, the amino transferase 5-aminolevulinic acid synthase catalyzes the addition of succinyl-CoA to glycine, the first step in heme biosynthesis. Other amino transferases participate in pathways important for neurological function and metabolism. For example, glutamine-phenylpyruvate amino transferase, also known as glutamine transaminase K (GTK), catalyzes several reactions with a pyridoxal phosphate cofactor. GTK catalyzes the reversible conversion of L-glutamine and phenylpyruvate to 2-oxoglutaramate and L-phenylalanine. Other

amino acid substrates for GTK include L-methionine, L-histidine, and L-tyrosine. GTK also catalyzes the conversion of kynurenine to kynurenic acid, a tryptophan metabolite that is an antagonist of the N-methyl-D-aspartate (NMDA) receptor in the brain and may exert a neuromodulatory function. Alteration of the kynurenine metabolic pathway may be associated with several neurological disorders. GTK also plays a role in the metabolism of halogenated xenobiotics conjugated to glutathione, leading to nephrotoxicity in rats and neurotoxicity in humans. GTK is expressed in kidney, liver, and brain. Both human and rat GTKs contain a putative pyridoxal phosphate binding site (ExpASY ENZYME: EC 2.6.1.64; Perry, S.J. et al. (1993) *Mol. Pharmacol.* 43:660-665; Perry, S. et al. (1995) *FEBS Lett.* 360:277-280; and Alberati-Giani, D. et al. (1995) *J. Neurochem.* 64:1448-1455). A second amino transferase associated with this pathway is kynurenine/ $\alpha$ -amino adipate amino transferase (AadAT). AadAT catalyzes the reversible conversion of  $\alpha$ -amino adipate and  $\alpha$ -ketoglutarate to  $\alpha$ -ketoadipate and L-glutamate during lysine metabolism. AadAT also catalyzes the transamination of kynurenine to kynurenic acid. A cytosolic AadAT is expressed in rat kidney, liver, and brain (Nakatani, Y. et al. (1970) *Biochim. Biophys. Acta* 198:219-228; Buchli, R. et al. (1995) *J. Biol. Chem.* 270:29330-29335).

Glycosyl transferases include the mammalian UDP-glucouronosyl transferases, a family of membrane-bound microsomal enzymes catalyzing the transfer of glucouronic acid to lipophilic substrates in reactions that play important roles in detoxification and excretion of drugs, carcinogens, and other foreign substances. Another mammalian glycosyl transferase, mammalian UDP-galactose-ceramide galactosyl transferase, catalyzes the transfer of galactose to ceramide in the synthesis of galactocerebrosides in myelin membranes of the nervous system. The UDP-glycosyl transferases share a conserved signature domain of about 50 amino acid residues (PROSITE: PDOC00359, <http://expasy.hcuge.ch/sprot/prosite.html>).

Methyl transferases are involved in a variety of pharmacologically important processes. Nicotinamide N-methyl transferase catalyzes the N-methylation of nicotinamides and other pyridines, an important step in the cellular handling of drugs and other foreign compounds. Phenylethanolamine N-methyl transferase catalyzes the conversion of noradrenalin to adrenalin. 6-O-methylguanine-DNA methyl transferase reverses DNA methylation, an important step in carcinogenesis. Uroporphyrin-III C-methyl transferase, which catalyzes the transfer of two methyl groups from S-adenosyl-L-methionine to uroporphyrinogen III, is the first specific enzyme in the biosynthesis of cobalamin, a dietary enzyme whose uptake is deficient in pernicious anemia. Protein-arginine methyl transferases catalyze the posttranslational methylation of arginine residues in proteins, resulting in the mono- and dimethylation of arginine on the guanidino group. Substrates include histones, myelin basic protein, and heterogeneous nuclear ribonucleoproteins involved in mRNA processing, splicing, and transport. Protein-arginine methyl transferase interacts with proteins upregulated by mitogens, with proteins involved in chronic lymphocytic leukemia, and with



interferon, suggesting an important role for methylation in cytokine receptor signaling (Lin, W.-J. et al. (1996) *J. Biol. Chem.* 271:15034-15044; Abramovich, C. et al. (1997) *EMBO J.* 16:260-266; and Scott, H.S. et al. (1998) *Genomics* 48:330-340).

Phosphotransferases catalyze the transfer of high-energy phosphate groups and are important in energy-requiring and -releasing reactions. The metabolic enzyme creatine kinase catalyzes the reversible phosphate transfer between creatine/creatine phosphate and ATP/ADP. Glycocyamine kinase catalyzes phosphate transfer from ATP to guanidoacetate, and arginine kinase catalyzes phosphate transfer from ATP to arginine. A cysteine-containing active site is conserved in this family (PROSITE: PDOC00103).

Prenyl transferases are heterodimers, consisting of an alpha and a beta subunit, that catalyze the transfer of an isoprenyl group. An example of a prenyl transferase is the mammalian protein farnesyl transferase. The alpha subunit of farnesyl transferase consists of 5 repeats of 34 amino acids each, with each repeat containing an invariant tryptophan (PROSITE: PDOC00703).

Saccharyl transferases are glycosylating enzymes involved in a variety of metabolic processes.

Oligosaccharyl transferase-48, for example, is a receptor for advanced glycation endproducts. Accumulation of these endproducts is observed in vascular complications of diabetes, macrovascular disease, renal insufficiency, and Alzheimer's disease (Thornalley, P.J. (1998) *Cell Mol. Biol. (Noisy-Le-Grand)* 44:1013-1023).

Coenzyme A (CoA) transferase catalyzes the transfer of CoA between two carboxylic acids.

Succinyl CoA:3-oxoacid CoA transferase, for example, transfers CoA from succinyl-CoA to a recipient such as acetoacetate. Acetoacetate is essential to the metabolism of ketone bodies, which accumulate in tissues affected by metabolic disorders such as diabetes (PROSITE: PDOC00980).

### Hydrolases

Hydrolysis is the breaking of a covalent bond in a substrate by introduction of a molecule of water. The reaction involves a nucleophilic attack by the water molecule's oxygen atom on a target bond in the substrate. The water molecule is split across the target bond, breaking the bond and generating two product molecules. Hydrolases participate in reactions essential to such functions as synthesis and degradation of cell components, and for regulation of cell functions including cell signaling, cell proliferation, inflammation, apoptosis, secretion and excretion. Hydrolases are involved in key steps in disease processes involving these functions. Hydrolytic enzymes, or hydrolases, may be grouped by substrate specificity into classes including phosphatases, peptidases, lysophospholipases, phosphodiesterases, glycosidases, and glyoxalases.

Phosphatases hydrolytically remove phosphate groups from proteins, an energy-providing step that regulates many cellular processes, including intracellular signaling pathways that in turn control cell growth and differentiation, cell-cell contact, the cell cycle, and oncogenesis.

Lysophospholipases (LPLs) regulate intracellular lipids by catalyzing the hydrolysis of ester

bonds to remove an acyl group, a key step in lipid degradation. Small LPL isoforms, approximately 15-30 kD, function as hydrolases; larger isoforms function both as hydrolases and transacylases. A particular substrate for LPLs, lysophosphatidylcholine, causes lysis of cell membranes. LPL activity is regulated by signaling molecules important in numerous pathways, including the inflammatory response.

Peptidases, also called proteases, cleave peptide bonds that form the backbone of peptide or protein chains. Proteolytic processing is essential to cell growth, differentiation, remodeling, and homeostasis as well as inflammation and immune response. Since typical protein half-lives range from hours to a few days, peptidases are continually cleaving precursor proteins to their active form, removing signal sequences from targeted proteins, and degrading aged or defective proteins. Peptidases function in bacterial, parasitic, and viral invasion and replication within a host. Examples of peptidases include trypsin and chymotrypsin (components of the complement cascade and the blood-clotting cascade) lysosomal cathepsins, calpains, pepsin, renin, and chymosin (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York NY, pp. 1-5).

The phosphodiesterases catalyze the hydrolysis of one of the two ester bonds in a phosphodiester compound. Phosphodiesterases are therefore crucial to a variety of cellular processes. Phosphodiesterases include DNA and RNA endo- and exo-nucleases, which are essential to cell growth and replication as well as protein synthesis. Another phosphodiesterase is acid sphingomyelinase, which hydrolyzes the membrane phospholipid sphingomyelin to ceramide and phosphorylcholine. Phosphorylcholine is used in the synthesis of phosphatidylcholine, which is involved in numerous intracellular signaling pathways. Ceramide is an essential precursor for the generation of gangliosides, membrane lipids found in high concentration in neural tissue. Defective acid sphingomyelinase phosphodiesterase leads to a build-up of sphingomyelin molecules in lysosomes, resulting in Niemann-Pick disease.

Glycosidases catalyze the cleavage of hemiacetyl bonds of glycosides, which are compounds that contain one or more sugar. Mammalian lactase-phlorizin hydrolase, for example, is an intestinal enzyme that splits lactose. Mammalian beta-galactosidase removes the terminal galactose from gangliosides, glycoproteins, and glycosaminoglycans, and deficiency of this enzyme is associated with a gangliosidosis known as Morquio disease type B. Vertebrate lysosomal alpha-glucosidase, which hydrolyzes glycogen, maltose, and isomaltose, and vertebrate intestinal sucrase-isomaltase, which hydrolyzes sucrose, maltose, and isomaltose, are widely distributed members of this family with highly conserved sequences at their active sites.

The glyoxylase system is involved in gluconeogenesis, the production of glucose from storage compounds in the body. It consists of glyoxylase I, which catalyzes the formation of S-D-lactoylglutathione from methyglyoxal, a side product of triose-phosphate energy metabolism, and

glyoxylase II, which hydrolyzes S-D-lactoylglutathione to D-lactic acid and reduced glutathione. Glyoxylases are involved in hyperglycemia, non-insulin-dependent diabetes mellitus, the detoxification of bacterial toxins, and in the control of cell proliferation and microtubule assembly.

### Lyases

5 Lyases are a class of enzymes that catalyze the cleavage of C-C, C-O, C-N, C-S, C-(halide), P-O or other bonds without hydrolysis or oxidation to form two molecules, at least one of which contains a double bond (Stryer, L. (1995) Biochemistry W.H. Freeman and Co. New York, NY p.620). Lyases are critical components of cellular biochemistry with roles in metabolic energy production including fatty acid metabolism, as well as other diverse enzymatic processes. Further  
10 classification of lyases reflects the type of bond cleaved as well as the nature of the cleaved group.

The group of C-C lyases include carboxyl-lyases (decarboxylases), aldehyde-lyases (aldolases), oxo-acid-lyases and others. The C-O lyase group includes hydro-lyases, lyases acting on polysaccharides and other lyases. The C-N lyase group includes ammonia-lyases, amidine-lyases, amine-lyases (deaminases) and other lyases.

15 Proper regulation of lyases is critical to normal physiology. For example, mutation induced deficiencies in the uroporphyrinogen decarboxylase can lead to photosensitive cutaneous lesions in the genetically-linked disorder familial porphyria cutanea tarda (Mendez, M. et al. (1998) *Am. J. Genet.* 63:1363-1375). It has also been shown that adenosine deaminase (ADA) deficiency stems from genetic mutations in the ADA gene, resulting in the disorder severe combined  
20 immunodeficiency disease (SCID) (Hershfield, M.S. (1998) *Semin. Hematol.* 35:291-298).

### Isomerases

Isomerases are a class of enzymes that catalyze geometric or structural changes within a molecule to form a single product. This class includes racemases and epimerases, cis-trans-isomerases, intramolecular oxidoreductases, intramolecular transferases (mutases) and intramolecular  
25 lyases. Isomerases are critical components of cellular biochemistry with roles in metabolic energy production including glycolysis, as well as other diverse enzymatic processes (Stryer, L. (1995) Biochemistry, W.H. Freeman and Co., New York NY, pp.483-507).

Racemases are a subset of isomerases that catalyze inversion of a molecule's configuration around the asymmetric carbon atom in a substrate having a single center of asymmetry, thereby  
30 interconverting two racemers. Epimerases are another subset of isomerases that catalyze inversion of configuration around an asymmetric carbon atom in a substrate with more than one center of symmetry, thereby interconverting two epimers. Racemases and epimerases can act on amino acids and derivatives, hydroxy acids and derivatives, as well as carbohydrates and derivatives. The interconversion of UDP-galactose and UDP-glucose is catalyzed by UDP-galactose-4'-epimerase.  
35 Proper regulation and function of this epimerase is essential to the synthesis of glycoproteins and glycolipids. Elevated blood galactose levels have been correlated with UDP-galactose-4'-epimerase

deficiency in screening programs of infants (Gitzelmann, R. (1972) *Helv. Paediat. Acta* 27:125-130).

Oxidoreductases can be isomerases as well. Oxidoreductases catalyze the reversible transfer of electrons from a substrate that becomes oxidized to a substrate that becomes reduced. This class of enzymes includes dehydrogenases, hydroxylases, oxidases, oxygenases, peroxidases, and  
5 reductases. Proper maintenance of oxidoreductase levels is physiologically important. For example, genetically-linked deficiencies in lipoamide dehydrogenase can result in lactic acidosis (Robinson, B.H. et al. (1977) *Pediat. Res.* 11:1198-1202).

Another subgroup of isomerases are the transferases (or mutases). Transferases transfer a chemical group from one compound (the donor) to another compound (the acceptor). The types of  
10 groups transferred by these enzymes include acyl groups, amino groups, phosphate groups (phosphotransferases or phosphomutases), and others. The transferase carnitine palmitoyltransferase is an important component of fatty acid metabolism. Genetically-linked deficiencies in this transferase can lead to myopathy (Scriver, C.R. et al. (1995) The Metabolic and Molecular Basis of Inherited Disease, McGraw-Hill, New York NY, pp.1501-1533).

15 Yet another subgroup of isomerases are the topoisomerases. Topoisomerases are enzymes that affect the topological state of DNA. For example, defects in topoisomerases or their regulation can affect normal physiology. Reduced levels of topoisomerase II have been correlated with some of the DNA processing defects associated with the disorder ataxia-telangiectasia (Singh, S.P. et al. (1988) *Nucleic Acids Res.* 16:3919-3929).

## 20 Ligases

Ligases catalyze the formation of a bond between two substrate molecules. The process involves the hydrolysis of a pyrophosphate bond in ATP or a similar energy donor. Ligases are classified based on the nature of the type of bond they form, which can include carbon-oxygen, carbon-sulfur, carbon-nitrogen, carbon-carbon and phosphoric ester bonds.

25 Ligases forming carbon-oxygen bonds include the aminoacyl-transfer RNA (tRNA) synthetases which are important RNA-associated enzymes with roles in translation. Protein biosynthesis depends on each amino acid forming a linkage with the appropriate tRNA. The aminoacyl-tRNA synthetases are responsible for the activation and correct attachment of an amino acid with its cognate tRNA. The 20 aminoacyl-tRNA synthetase enzymes can be divided into two  
30 structural classes, and each class is characterized by a distinctive topology of the catalytic domain. Class I enzymes contain a catalytic domain based on the nucleotide-binding Rossmann fold. Class II enzymes contain a central catalytic domain, which consists of a seven-stranded antiparallel  $\beta$ -sheet motif, as well as N- and C- terminal regulatory domains. Class II enzymes are separated into two groups based on the heterodimeric or homodimeric structure of the enzyme; the latter group is further  
35 subdivided by the structure of the N- and C-terminal regulatory domains (Hartlein, M. and S. Cusack (1995) *J. Mol. Evol.* 40:519-530). Autoantibodies against aminoacyl-tRNAs are generated by

patients with dermatomyositis and polymyositis, and correlate strongly with complicating interstitial lung disease (ILD). These antibodies appear to be generated in response to viral infection, and coxsackie virus has been used to induce experimental viral myositis in animals.

Ligases forming carbon-sulfur bonds (Acid-thiol ligases) mediate a large number of cellular biosynthetic intermediary metabolism processes involve intermolecular transfer of carbon atom-containing substrates (carbon substrates). Examples of such reactions include the tricarboxylic acid cycle, synthesis of fatty acids and long-chain phospholipids, synthesis of alcohols and aldehydes, synthesis of intermediary metabolites, and reactions involved in the amino acid degradation pathways. Some of these reactions require input of energy, usually in the form of conversion of ATP to either ADP or AMP and pyrophosphate.

In many cases, a carbon substrate is derived from a small molecule containing at least two carbon atoms. The carbon substrate is often covalently bound to a larger molecule which acts as a carbon substrate carrier molecule within the cell. In the biosynthetic mechanisms described above, the carrier molecule is coenzyme A. Coenzyme A (CoA) is structurally related to derivatives of the nucleotide ADP and consists of 4'-phosphopantetheine linked via a phosphodiester bond to the alpha phosphate group of adenosine 3',5'-bisphosphate. The terminal thiol group of 4'-phosphopantetheine acts as the site for carbon substrate bond formation. The predominant carbon substrates which utilize CoA as a carrier molecule during biosynthesis and intermediary metabolism in the cell are acetyl, succinyl, and propionyl moieties, collectively referred to as acyl groups. Other carbon substrates include enoyl lipid, which acts as a fatty acid oxidation intermediate, and carnitine, which acts as an acetyl-CoA flux regulator/ mitochondrial acyl group transfer protein. Acyl-CoA and acetyl-CoA are synthesized in the cell by acyl-CoA synthetase and acetyl-CoA synthetase, respectively.

Activation of fatty acids is mediated by at least three forms of acyl-CoA synthetase activity: i) acetyl-CoA synthetase, which activates acetate and several other low molecular weight carboxylic acids and is found in muscle mitochondria and the cytosol of other tissues; ii) medium-chain acyl-CoA synthetase, which activates fatty acids containing between four and eleven carbon atoms (predominantly from dietary sources), and is present only in liver mitochondria; and iii) acyl CoA synthetase, which is specific for long chain fatty acids with between six and twenty carbon atoms, and is found in microsomes and the mitochondria. Proteins associated with acyl-CoA synthetase activity have been identified from many sources including bacteria, yeast, plants, mouse, and man. The activity of acyl-CoA synthetase may be modulated by phosphorylation of the enzyme by cAMP-dependent protein kinase.

Ligases forming carbon-nitrogen bonds include amide synthases such as glutamine synthetase (glutamate-ammonia ligase) that catalyzes the amination of glutamic acid to glutamine by ammonia using the energy of ATP hydrolysis. Glutamine is the primary source for the amino group in various amide transfer reactions involved in de novo pyrimidine nucleotide synthesis and in purine

and pyrimidine ribonucleotide interconversions. Overexpression of glutamine synthetase has been observed in primary liver cancer (Christa, L. et al. (1994) *Gastroent.* 106:1312-1320).

Acid-amino-acid ligases (peptide synthases) are represented by the ubiquitin proteases which are associated with the ubiquitin conjugation system (UCS), a major pathway for the degradation of cellular proteins in eukaryotic cells and some bacteria. The UCS mediates the elimination of abnormal proteins and regulates the half-lives of important regulatory proteins that control cellular processes such as gene transcription and cell cycle progression. In the UCS pathway, proteins targeted for degradation are conjugated to a ubiquitin (Ub), a small heat stable protein. Ub is first activated by a ubiquitin-activating enzyme (E1), and then transferred to one of several Ub-conjugating enzymes (E2). E2 then links the Ub molecule through its C-terminal glycine to an internal lysine (acceptor lysine) of a target protein. The ubiquitinated protein is then recognized and degraded by proteasome, a large, multisubunit proteolytic enzyme complex, and ubiquitin is released for reutilization by ubiquitin protease. The UCS is implicated in the degradation of mitotic cyclic kinases, oncoproteins, tumor suppressor genes such as p53, viral proteins, cell surface receptors associated with signal transduction, transcriptional regulators, and mutated or damaged proteins (Ciechanover, A. (1994) *Cell* 79:13-21). A murine proto-oncogene, *Unp*, encodes a nuclear ubiquitin protease whose overexpression leads to oncogenic transformation of NIH3T3 cells, and the human homolog of this gene is consistently elevated in small cell tumors and adenocarcinomas of the lung (Gray, D.A. (1995) *Oncogene* 10:2179-2183).

Cyclo-ligases and other carbon-nitrogen ligases comprise various enzymes and enzyme complexes that participate in the de novo pathways to purine and pyrimidine biosynthesis. Because these pathways are critical to the synthesis of nucleotides for replication of both RNA and DNA, many of these enzymes have been the targets of clinical agents for the treatment of cell proliferative disorders such as cancer and infectious diseases.

Purine biosynthesis occurs de novo from the amino acids glycine and glutamine, and other small molecules. Three of the key reactions in this process are catalyzed by a trifunctional enzyme composed of glycinamide-ribonucleotide synthetase (GARS), aminoimidazole ribonucleotide synthetase (AIRS), and glycinamide ribonucleotide transformylase (GART). Together these three enzymes combine ribosylamine phosphate with glycine to yield phosphoribosyl aminoimidazole, a precursor to both adenylylate and guanylylate nucleotides. This trifunctional protein has been implicated in the pathology of Downs syndrome (Aimi, J. et al. (1990) *Nucleic Acid Res.* 18:6665-6672). Adenylosuccinate synthetase catalyzes a later step in purine biosynthesis that converts inosinic acid to adenylosuccinate, a key step on the path to ATP synthesis. This enzyme is also similar to another carbon-nitrogen ligase, argininosuccinate synthetase, that catalyzes a similar reaction in the urea cycle (Powell, S.M. et al. (1992) *FEBS Lett.* 303:4-10).

Like the de novo biosynthesis of purines, de novo synthesis of the pyrimidine nucleotides

uridylylate and cytidylylate also arises from a common precursor, in this instance the nucleotide orotidylylate derived from orotate and phosphoribosyl pyrophosphate (PPRP). Again a trifunctional enzyme comprising three carbon-nitrogen ligases plays a key role in the process. In this case the enzymes aspartate transcarbamylase (ATCase), carbamyl phosphate synthetase II, and dihydroorotase (DHOase) are encoded by a single gene called CAD. Together these three enzymes combine the initial reactants in pyrimidine biosynthesis, glutamine, CO<sub>2</sub>, and ATP to form dihydroorotate, the precursor to orotate and orotidylylate (Iwahana, H. et al. (1996) *Biochem. Biophys. Res. Commun.* 219:249-255). Further steps then lead to the synthesis of uridine nucleotides from orotidylylate. Cytidine nucleotides are derived from uridine-5'-triphosphate (UTP) by the amidation of UTP using glutamine as the amino donor and the enzyme CTP synthetase. Regulatory mutations in the human CTP synthetase are believed to confer multi-drug resistance to agents widely used in cancer therapy (Yamauchi, M. et al. (1990) *EMBO J.* 9:2095-2099).

Ligases forming carbon-carbon bonds include the carboxylases acetyl-CoA carboxylase and pyruvate carboxylase. Acetyl-CoA carboxylase catalyzes the carboxylation of acetyl-CoA from CO<sub>2</sub> and H<sub>2</sub>O using the energy of ATP hydrolysis. Acetyl-CoA carboxylase is the rate-limiting step in the biogenesis of long-chain fatty acids. Two isoforms of acetyl-CoA carboxylase, types I and types II, are expressed in human in a tissue-specific manner (Ha, J. et al. (1994) *Eur. J. Biochem.* 219:297-306). Pyruvate carboxylase is a nuclear-encoded mitochondrial enzyme that catalyzes the conversion of pyruvate to oxaloacetate, a key intermediate in the citric acid cycle.

Ligases forming phosphoric ester bonds include the DNA ligases involved in both DNA replication and repair. DNA ligases seal phosphodiester bonds between two adjacent nucleotides in a DNA chain using the energy from ATP hydrolysis to first activate the free 5'-phosphate of one nucleotide and then react it with the 3'-OH group of the adjacent nucleotide. This resealing reaction is used in both DNA replication to join small DNA fragments called Okazaki fragments that are transiently formed in the process of replicating new DNA, and in DNA repair. DNA repair is the process by which accidental base changes, such as those produced by oxidative damage, hydrolytic attack, or uncontrolled methylation of DNA, are corrected before replication or transcription of the DNA can occur. Bloom's syndrome is an inherited human disease in which individuals are partially deficient in DNA ligation and consequently have an increased incidence of cancer (Alberts, B. et al. (1994) *The Molecular Biology of the Cell*, Garland Publishing Inc., New York NY, p. 247).

### **Molecules Associated with Growth and Development**

Human growth and development requires the spatial and temporal regulation of cell differentiation, cell proliferation, and apoptosis. These processes coordinately control reproduction, aging, embryogenesis, morphogenesis, organogenesis, and tissue repair and maintenance. At the cellular level, growth and development is governed by the cell's decision to enter into or exit from

the cell division cycle and by the cell's commitment to a terminally differentiated state. These decisions are made by the cell in response to extracellular signals and other environmental cues it receives. The following discussion focuses on the molecular mechanisms of cell division, reproduction, cell differentiation and proliferation, apoptosis, and aging.

## 5 Cell Division

Cell division is the fundamental process by which all living things grow and reproduce. In unicellular organisms such as yeast and bacteria, each cell division doubles the number of organisms, while in multicellular species many rounds of cell division are required to replace cells lost by wear or by programmed cell death, and for cell differentiation to produce a new tissue or organ. Details of  
10 the cell division cycle may vary, but the basic process consists of three principle events. The first event, interphase, involves preparations for cell division, replication of the DNA, and production of essential proteins. In the second event, mitosis, the nuclear material is divided and separates to opposite sides of the cell. The final event, cytokinesis, is division and fission of the cell cytoplasm. The sequence and timing of cell cycle transitions is under the control of the cell cycle regulation  
15 system which controls the process by positive or negative regulatory circuits at various check points.

Regulated progression of the cell cycle depends on the integration of growth control pathways with the basic cell cycle machinery. Cell cycle regulators have been identified by selecting for human and yeast cDNAs that block or activate cell cycle arrest signals in the yeast mating pheromone pathway when they are overexpressed. Known regulators include human CPR (cell cycle  
20 progression restoration) genes, such as CPR8 and CPR2, and yeast CDC (cell division control) genes, including CDC91, that block the arrest signals. The CPR genes express a variety of proteins including cyclins, tumor suppressor binding proteins, chaperones, transcription factors, translation factors, and RNA-binding proteins (Edwards, M.C. et al.(1997) Genetics 147:1063-1076).

Several cell cycle transitions, including the entry and exit of a cell from mitosis, are  
25 dependent upon the activation and inhibition of cyclin-dependent kinases (Cdks). The Cdks are composed of a kinase subunit, Cdk, and an activating subunit, cyclin, in a complex that is subject to many levels of regulation. There appears to be a single Cdk in Saccharomyces cerevisiae and Saccharomyces pombe whereas mammals have a variety of specialized Cdks. Cyclins act by binding to and activating cyclin-dependent protein kinases which then phosphorylate and activate selected  
30 proteins involved in the mitotic process. The Cdk-cyclin complex is both positively and negatively regulated by phosphorylation, and by targeted degradation involving molecules such as CDC4 and CDC53. In addition, Cdks are further regulated by binding to inhibitors and other proteins such as Suc1 that modify their specificity or accessibility to regulators (Patra, D. and W.G. Dunphy (1996) Genes Dev. 10:1503-1515; and Mathias, N. et al. (1996) Mol. Cell Biol. 16:6634-6643).

## 35 Reproduction

The male and female reproductive systems are complex and involve many aspects of growth



and development. The anatomy and physiology of the male and female reproductive systems are reviewed in (Guyton, A.C. (1991) Textbook of Medical Physiology, W.B. Saunders Co., Philadelphia PA, pp. 899-928).

5 The male reproductive system includes the process of spermatogenesis, in which the sperm are formed, and male reproductive functions are regulated by various hormones and their effects on accessory sexual organs, cellular metabolism, growth, and other bodily functions.

Spermatogenesis begins at puberty as a result of stimulation by gonadotropic hormones released from the anterior pituitary. Immature sperm (spermatogonia) undergo several mitotic cell divisions before undergoing meiosis and full maturation. The testes secrete several male sex  
10 hormones, the most abundant being testosterone, that is essential for growth and division of the immature sperm, and for the masculine characteristics of the male body. Three other male sex hormones, gonadotropin-releasing hormone (GnRH), luteinizing hormone (LH), and follicle-stimulating hormone (FSH) control sexual function.

The uterus, ovaries, fallopian tubes, vagina, and breasts comprise the female reproductive  
15 system. The ovaries and uterus are the source of ova and the location of fetal development, respectively. The fallopian tubes and vagina are accessory organs attached to the top and bottom of the uterus, respectively. Both the uterus and ovaries have additional roles in the development and loss of reproductive capability during a female's lifetime. The primary role of the breasts is lactation. Multiple endocrine signals from the ovaries, uterus, pituitary, hypothalamus, adrenal glands, and  
20 other tissues coordinate reproduction and lactation. These signals vary during the monthly menstruation cycle and during the female's lifetime. Similarly, the sensitivity of reproductive organs to these endocrine signals varies during the female's lifetime.

A combination of positive and negative feedback to the ovaries, pituitary and hypothalamus glands controls physiologic changes during the monthly ovulation and endometrial cycles. The  
25 anterior pituitary secretes two major gonadotropin hormones, follicle-stimulating hormone (FSH) and luteinizing hormone (LH), regulated by negative feedback of steroids, most notably by ovarian estradiol. If fertilization does not occur, estrogen and progesterone levels decrease. This sudden reduction of the ovarian hormones leads to menstruation, the desquamation of the endometrium.

Hormones further govern all the steps of pregnancy, parturition, lactation, and menopause.  
30 During pregnancy large quantities of human chorionic gonadotropin (hCG), estrogens, progesterone, and human chorionic somatomammotropin (hCS) are formed by the placenta. hCG, a glycoprotein similar to luteinizing hormone, stimulates the corpus luteum to continue producing more progesterone and estrogens, rather than to involute as occurs if the ovum is not fertilized. hCS is similar to growth hormone and is crucial for fetal nutrition.

35 The female breast also matures during pregnancy. Large amounts of estrogen secreted by the placenta trigger growth and branching of the breast milk ductal system while lactation is initiated by

the secretion of prolactin by the pituitary gland.

Parturition involves several hormonal changes that increase uterine contractility toward the end of pregnancy, as follows. The levels of estrogens increase more than those of progesterone.

Oxytocin is secreted by the neurohypophysis. Concomitantly, uterine sensitivity to oxytocin  
5 increases. The fetus itself secretes oxytocin, cortisol (from adrenal glands), and prostaglandins.

Menopause occurs when most of the ovarian follicles have degenerated. The ovary then produces less estradiol, reducing the negative feedback on the pituitary and hypothalamus glands. Mean levels of circulating FSH and LH increase, even as ovulatory cycles continue. Therefore, the ovary is less responsive to gonadotropins, and there is an increase in the time between menstrual  
10 cycles. Consequently, menstrual bleeding ceases and reproductive capability ends.

#### Cell Differentiation and Proliferation

Tissue growth involves complex and ordered patterns of cell proliferation, cell differentiation, and apoptosis. Cell proliferation must be regulated to maintain both the number of cells and their spatial organization. This regulation depends upon the appropriate expression of  
15 proteins which control cell cycle progression in response to extracellular signals, such as growth factors and other mitogens, and intracellular cues, such as DNA damage or nutrient starvation. Molecules which directly or indirectly modulate cell cycle progression fall into several categories, including growth factors and their receptors, second messenger and signal transduction proteins, oncogene products, tumor-suppressor proteins, and mitosis-promoting factors.

Growth factors were originally described as serum factors required to promote cell  
20 proliferation. Most growth factors are large, secreted polypeptides that act on cells in their local environment. Growth factors bind to and activate specific cell surface receptors and initiate intracellular signal transduction cascades. Many growth factor receptors are classified as receptor tyrosine kinases which undergo autophosphorylation upon ligand binding. Autophosphorylation  
25 enables the receptor to interact with signal transduction proteins characterized by the presence of SH2 or SH3 domains (Src homology regions 2 or 3). These proteins then modulate the activity state of small G-proteins, such as Ras, Rab, and Rho, along with GTPase activating proteins (GAPs), guanine nucleotide releasing proteins (GNRPs), and other guanine nucleotide exchange factors. Small G proteins act as molecular switches that activate other downstream events, such as mitogen-activated  
30 protein kinase (MAP kinase) cascades. MAP kinases ultimately activate transcription of mitosis-promoting genes.

In addition to growth factors, small signaling peptides and hormones also influence cell proliferation. These molecules bind primarily to another class of receptor, the trimeric G-protein coupled receptor (GPCR), found predominantly on the surface of immune, neuronal and  
35 neuroendocrine cells. Upon ligand binding, the GPCR activates a trimeric G protein which in turn triggers increased levels of intracellular second messengers such as phospholipase C, Ca<sup>2+</sup>, and

cyclic AMP. Most GPCR-mediated signaling pathways indirectly promote cell proliferation by causing the secretion or breakdown of other signaling molecules that have direct mitogenic effects. These signaling cascades often involve activation of kinases and phosphatases. Some growth factors, such as some members of the transforming growth factor beta (TGF- $\beta$ ) family, act on some cells to stimulate cell proliferation and on other cells to inhibit it. Growth factors may also stimulate a cell at one concentration and inhibit the same cell at another concentration. Most growth factors also have a multitude of other actions besides the regulation of cell growth and division: they can control the proliferation, survival, differentiation, migration, or function of cells depending on the circumstance. For example, the tumor necrosis factor/nerve growth factor (TNF/NGF) family can activate or inhibit cell death, as well as regulate proliferation and differentiation. The cell response depends on the type of cell, its stage of differentiation and transformation status, which surface receptors are stimulated, and the types of stimuli acting on the cell (Smith, A. et al. (1994) Cell 76:959-962; and Nocentini, G. et al. (1997) Proc. Natl. Acad. Sci. USA 94:6216-6221).

Neighboring cells in a tissue compete for growth factors, and when provided with "unlimited" quantities in a perfused system will grow to even higher cell densities before reaching density-dependent inhibition of cell division. Cells often demonstrate an anchorage dependence of cell division as well. This anchorage dependence may be associated with the formation of focal contacts linking the cytoskeleton with the extracellular matrix (ECM). The expression of ECM components can be stimulated by growth factors. For example, TGF- $\beta$  stimulates fibroblasts to produce a variety of ECM proteins, including fibronectin, collagen, and tenascin (Pearson, C.A. et al. (1988) EMBO J. 7:2677-2981). In fact, for some cell types specific ECM molecules, such as laminin or fibronectin, may act as growth factors. Tenascin-C and -R, expressed in developing and lesioned neural tissue, provide stimulatory/anti-adhesive or inhibitory properties, respectively, for axonal growth (Faissner, A. (1997) Cell Tissue Res. 290:331-341).

Cancers are associated with the activation of oncogenes which are derived from normal cellular genes. These oncogenes encode oncoproteins which convert normal cells into malignant cells. Some oncoproteins are mutant isoforms of the normal protein, and other oncoproteins are abnormally expressed with respect to location or amount of expression. The latter category of oncoprotein causes cancer by altering transcriptional control of cell proliferation. Five classes of oncoproteins are known to affect cell cycle controls. These classes include growth factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and cell-cycle control proteins. Viral oncogenes are integrated into the human genome after infection of human cells by certain viruses. Examples of viral oncogenes include v-src, v-abl, and v-fps.

Many oncogenes have been identified and characterized. These include sis, erbA, erbB, her-2, mutated G<sub>s</sub>, src, abl, ras, crk, jun, fos, myc, and mutated tumor-suppressor genes such as RB, p53, mdm2, Cip1, p16, and cyclin D. Transformation of normal genes to oncogenes may also occur by

chromosomal translocation. The Philadelphia chromosome, characteristic of chronic myeloid leukemia and a subset of acute lymphoblastic leukemias, results from a reciprocal translocation between chromosomes 9 and 22 that moves a truncated portion of the proto-oncogene *c-abl* to the breakpoint cluster region (*bcr*) on chromosome 22.

5 Tumor-suppressor genes are involved in regulating cell proliferation. Mutations which cause reduced or loss of function in tumor-suppressor genes result in uncontrolled cell proliferation. For example, the retinoblastoma gene product (RB), in a non-phosphorylated state, binds several early-response genes and suppresses their transcription, thus blocking cell division. Phosphorylation of RB causes it to dissociate from the genes, releasing the suppression, and allowing cell division to  
10 proceed.

#### Apoptosis

Apoptosis is the genetically controlled process by which unneeded or defective cells undergo programmed cell death. Selective elimination of cells is as important for morphogenesis and tissue remodeling as is cell proliferation and differentiation. Lack of apoptosis may result in hyperplasia  
15 and other disorders associated with increased cell proliferation. Apoptosis is also a critical component of the immune response. Immune cells such as cytotoxic T-cells and natural killer cells prevent the spread of disease by inducing apoptosis in tumor cells and virus-infected cells. In addition, immune cells that fail to distinguish self molecules from foreign molecules must be eliminated by apoptosis to avoid an autoimmune response.

20 Apoptotic cells undergo distinct morphological changes. Hallmarks of apoptosis include cell shrinkage, nuclear and cytoplasmic condensation, and alterations in plasma membrane topology. Biochemically, apoptotic cells are characterized by increased intracellular calcium concentration, fragmentation of chromosomal DNA, and expression of novel cell surface components.

The molecular mechanisms of apoptosis are highly conserved, and many of the key protein  
25 regulators and effectors of apoptosis have been identified. Apoptosis generally proceeds in response to a signal which is transduced intracellularly and results in altered patterns of gene expression and protein activity. Signaling molecules such as hormones and cytokines are known both to stimulate and to inhibit apoptosis through interactions with cell surface receptors. Transcription factors also play an important role in the onset of apoptosis. A number of downstream effector molecules,  
30 particularly proteases such as the cysteine proteases called caspases, have been implicated in the degradation of cellular components and the proteolytic activation of other apoptotic effectors.

#### Aging and Senescence

Studies of the aging process or senescence have shown a number of characteristic cellular and molecular changes (Fauci et al. (1998) Harrison's Principles of Internal Medicine, McGraw-Hill, New  
35 York NY, p.37). These characteristics include increases in chromosome structural abnormalities, DNA cross-linking, incidence of single-stranded breaks in DNA, losses in DNA methylation, and

degradation of telomere regions. In addition to these DNA changes, post-translational alterations of proteins increase including, deamidation, oxidation, cross-linking, and nonenzymatic glycation. Still further molecular changes occur in the mitochondria of aging cells through deterioration of structure. These changes eventually contribute to decreased function in every organ of the body.

5

### Biochemical Pathway Molecules

Biochemical pathways are responsible for regulating metabolism, growth and development, protein secretion and trafficking, environmental responses, and ecological interactions including immune response and response to parasites.

### 10 DNA replication

Deoxyribonucleic acid (DNA), the genetic material, is found in both the nucleus and mitochondria of human cells. The bulk of human DNA is nuclear, in the form of linear chromosomes, while mitochondrial DNA is circular. DNA replication begins at specific sites called origins of replication. Bidirectional synthesis occurs from the origin via two growing forks that move in opposite directions. Replication is semi-conservative, with each daughter duplex containing one old strand and its newly synthesized complementary partner. Proteins involved in DNA replication include DNA polymerases, DNA primase, telomerase, DNA helicase, topoisomerases, DNA ligases, replication factors, and DNA-binding proteins.

### DNA Recombination and Repair

20 Cells are constantly faced with replication errors and environmental assault (such as ultraviolet irradiation) that can produce DNA damage. Damage to DNA consists of any change that modifies the structure of the molecule. Changes to DNA can be divided into two general classes, single base changes and structural distortions. Any damage to DNA can produce a mutation, and the mutation may produce a disorder, such as cancer.

25 Changes in DNA are recognized by repair systems within the cell. These repair systems act to correct the damage and thus prevent any deleterious effects of a mutational event. Repair systems can be divided into three general types, direct repair, excision repair, and retrieval systems. Proteins involved in DNA repair include DNA polymerase, excision repair proteins, excision and cross link repair proteins, recombination and repair proteins, RAD51 proteins, and BLN and WRN proteins that are homologs of RecQ helicase. When the repair systems are eliminated, cells become exceedingly sensitive to environmental mutagens, such as ultraviolet irradiation. Patients with disorders associated with a loss in DNA repair systems often exhibit a high sensitivity to environmental mutagens. Examples of such disorders include xeroderma pigmentosum (XP), Bloom's syndrome (BS), and Werner's syndrome (WS) (Yamagata, K. et al. (1998) Proc. Natl. Acad. Sci. USA 95:8733-8738), ataxia telangiectasia, Cockayne's syndrome, and Fanconi's anemia.

35

Recombination is the process whereby new DNA sequences are generated by the movements

of large pieces of DNA. In homologous recombination, which occurs during meiosis and DNA repair, parent DNA duplexes align at regions of sequence similarity, and new DNA molecules form by the breakage and joining of homologous segments. Proteins involved include RAD51 recombinase. In site-specific recombination, two specific but not necessarily homologous DNA sequences are exchanged. In the immune system this process generates a diverse collection of antibody and T cell receptor genes. Proteins involved in site-specific recombination in the immune system include recombination activating genes 1 and 2 (RAG1 and RAG2). A defect in immune system site-specific recombination causes severe combined immunodeficiency disease in mice.

### RNA Metabolism

10 Ribonucleic acid (RNA) is a linear single-stranded polymer of four nucleotides, ATP, CTP, UTP, and GTP. In most organisms, RNA is transcribed as a copy of DNA, the genetic material of the organism. In retroviruses RNA rather than DNA serves as the genetic material. RNA copies of the genetic material encode proteins or serve various structural, catalytic, or regulatory roles in organisms. RNA is classified according to its cellular localization and function. Messenger RNAs (mRNAs) encode polypeptides. Ribosomal RNAs (rRNAs) are assembled, along with ribosomal proteins, into ribosomes, which are cytoplasmic particles that translate mRNA into polypeptides. Transfer RNAs (tRNAs) are cytosolic adaptor molecules that function in mRNA translation by recognizing both an mRNA codon and the amino acid that matches that codon. Heterogeneous nuclear RNAs (hnRNAs) include mRNA precursors and other nuclear RNAs of various sizes. Small nuclear RNAs (snRNAs) are a part of the nuclear spliceosome complex that removes intervening, non-coding sequences (introns) and rejoins exons in pre-mRNAs.

### RNA Transcription

The transcription process synthesizes an RNA copy of DNA. Proteins involved include multi-subunit RNA polymerases, transcription factors IIA, IIB, IID, IIE, IIF, IIH, and III. Many transcription factors incorporate DNA-binding structural motifs which comprise either  $\alpha$ -helices or  $\beta$ -sheets that bind to the major groove of DNA. Four well-characterized structural motifs are helix-turn-helix, zinc finger, leucine zipper, and helix-loop-helix.

### RNA Processing

Various proteins are necessary for processing of transcribed RNAs in the nucleus. Pre-mRNA processing steps include capping at the 5' end with methylguanosine, polyadenylating the 3' end, and splicing to remove introns. The spliceosomal complex is comprised of five small nuclear ribonucleoprotein particles (snRNPs) designated U1, U2, U4, U5, and U6. Each snRNP contains a single species of snRNA and about ten proteins. The RNA components of some snRNPs recognize and base-pair with intron consensus sequences. The protein components mediate spliceosome assembly and the splicing reaction. Autoantibodies to snRNP proteins are found in the blood of patients with systemic lupus erythematosus (Stryer, L. (1995) Biochemistry W.H. Freeman and

Company, New York NY, p. 863).

Heterogeneous nuclear ribonucleoproteins (hnRNPs) have been identified that have roles in splicing, exporting of the mature RNAs to the cytoplasm, and mRNA translation (Biamonti, G. et al. (1998) Clin. Exp. Rheumatol. 16:317-326). Some examples of hnRNPs include the yeast proteins

5 Hrp1p, involved in cleavage and polyadenylation at the 3' end of the RNA; Cbp80p, involved in capping the 5' end of the RNA; and Npl3p, a homolog of mammalian hnRNP A1, involved in export of mRNA from the nucleus (Shen, E.C. et al. (1998) Genes Dev. 12:679-691). HnRNPs have been shown to be important targets of the autoimmune response in rheumatic diseases (Biamonti, supra).

Many snRNP proteins, hnRNP proteins, and alternative splicing factors are characterized by

10 an RNA recognition motif (RRM). (Reviewed in Birney, E. et al. (1993) Nucleic Acids Res. 21:5803-5816.) The RRM is about 80 amino acids in length and forms four  $\beta$ -strands and two  $\alpha$ -helices arranged in an  $\alpha/\beta$  sandwich. The RRM contains a core RNP-1 octapeptide motif along with surrounding conserved sequences.

#### RNA Stability and Degradation

15 RNA helicases alter and regulate RNA conformation and secondary structure by using energy derived from ATP hydrolysis to destabilize and unwind RNA duplexes. The most well-characterized and ubiquitous family of RNA helicases is the DEAD-box family, so named for the conserved B-type ATP-binding motif which is diagnostic of proteins in this family. Over 40 DEAD-box helicases have been identified in organisms as diverse as bacteria, insects, yeast, amphibians, mammals, and plants.

20 DEAD-box helicases function in diverse processes such as translation initiation, splicing, ribosome assembly, and RNA editing, transport, and stability. Some DEAD-box helicases play tissue- and stage-specific roles in spermatogenesis and embryogenesis. (Reviewed in Linder, P. et al. (1989) Nature 337:121-122.)

Overexpression of the DEAD-box 1 protein (DDX1) may play a role in the progression of

25 neuroblastoma (Nb) and retinoblastoma (Rb) tumors. Other DEAD-box helicases have been implicated either directly or indirectly in ultraviolet light-induced tumors, B cell lymphoma, and myeloid malignancies. (Reviewed in Godbout, R. et al. (1998) J. Biol. Chem. 273:21161-21168.)

Ribonucleases (RNases) catalyze the hydrolysis of phosphodiester bonds in RNA chains, thus cleaving the RNA. For example, RNase P is a ribonucleoprotein enzyme which cleaves the 5' end of

30 pre-tRNAs as part of their maturation process. RNase H digests the RNA strand of an RNA/DNA hybrid. Such hybrids occur in cells invaded by retroviruses, and RNase H is an important enzyme in the retroviral replication cycle. RNase H domains are often found as a domain associated with reverse transcriptases. RNase activity in serum and cell extracts is elevated in a variety of cancers and infectious diseases (Schein, C.H. (1997) Nat. Biotechnol. 15:529-536). Regulation of RNase

35 activity is being investigated as a means to control tumor angiogenesis, allergic reactions, viral infection and replication, and fungal infections.

### Protein Translation

The eukaryotic ribosome is composed of a 60S (large) subunit and a 40S (small) subunit, which together form the 80S ribosome. In addition to the 18S, 28S, 5S, and 5.8S rRNAs, the ribosome also contains more than fifty proteins. The ribosomal proteins have a prefix which denotes the subunit to which they belong, either L (large) or S (small). Three important sites are identified on the ribosome. The aminoacyl-tRNA site (A site) is where charged tRNAs (with the exception of the initiator-tRNA) bind on arrival at the ribosome. The peptidyl-tRNA site (P site) is where new peptide bonds are formed, as well as where the initiator tRNA binds. The exit site (E site) is where deacylated tRNAs bind prior to their release from the ribosome. (Translation is reviewed in Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, pp. 875-908; and Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 119-138.)

### tRNA Charging

Protein biosynthesis depends on each amino acid forming a linkage with the appropriate tRNA. The aminoacyl-tRNA synthetases are responsible for the activation and correct attachment of an amino acid with its cognate tRNA. The 20 aminoacyl-tRNA synthetase enzymes can be divided into two structural classes, Class I and Class II. Autoantibodies against aminoacyl-tRNAs are generated by patients with dermatomyositis and polymyositis, and correlate strongly with complicating interstitial lung disease (ILD). These antibodies appear to be generated in response to viral infection, and coxsackie virus has been used to induce experimental viral myositis in animals.

### Translation Initiation

Initiation of translation can be divided into three stages. The first stage brings an initiator transfer RNA (Met-tRNA<sub>i</sub>) together with the 40S ribosomal subunit to form the 43S preinitiation complex. The second stage binds the 43S preinitiation complex to the mRNA, followed by migration of the complex to the correct AUG initiation codon. The third stage brings the 60S ribosomal subunit to the 40S subunit to generate an 80S ribosome at the initiation codon. Regulation of translation primarily involves the first and second stage in the initiation process (Pain, V.M. (1996) Eur. J. Biochem. 236:747-771).

Several initiation factors, many of which contain multiple subunits, are involved in bringing an initiator tRNA and 40S ribosomal subunit together. eIF2, a guanine nucleotide binding protein, recruits the initiator tRNA to the 40S ribosomal subunit. Only when eIF2 is bound to GTP does it associate with the initiator tRNA. eIF2B, a guanine nucleotide exchange protein, is responsible for converting eIF2 from the GDP-bound inactive form to the GTP-bound active form. Two other factors, eIF1A and eIF3 bind and stabilize the 40S subunit by interacting with 18S ribosomal RNA and specific ribosomal structural proteins. eIF3 is also involved in association of the 40S ribosomal subunit with mRNA. The Met-tRNA<sub>i</sub>, eIF1A, eIF3, and 40S ribosomal subunit together make up the 43S preinitiation complex (Pain, supra).



Additional factors are required for binding of the 43S preinitiation complex to an mRNA molecule, and the process is regulated at several levels. eIF4F is a complex consisting of three proteins: eIF4E, eIF4A, and eIF4G. eIF4E recognizes and binds to the mRNA 5'-terminal m<sup>7</sup>GTP cap, eIF4A is a bidirectional RNA-dependent helicase, and eIF4G is a scaffolding polypeptide.

5 eIF4G has three binding domains. The N-terminal third of eIF4G interacts with eIF4E, the central third interacts with eIF4A, and the C-terminal third interacts with eIF3 bound to the 43S preinitiation complex. Thus, eIF4G acts as a bridge between the 40S ribosomal subunit and the mRNA (Hentze, M.W. (1997) Science 275:500-501).

The ability of eIF4F to initiate binding of the 43S preinitiation complex is regulated by structural features of the mRNA. The mRNA molecule has an untranslated region (UTR) between the 5' cap and the AUG start codon. In some mRNAs this region forms secondary structures that impede binding of the 43S preinitiation complex. The helicase activity of eIF4A is thought to function in removing this secondary structure to facilitate binding of the 43S preinitiation complex (Pain, *supra*).

#### 15 Translation Elongation

Elongation is the process whereby additional amino acids are joined to the initiator methionine to form the complete polypeptide chain. The elongation factors EF1 $\alpha$ , EF1 $\beta$   $\gamma$ , and EF2 are involved in elongating the polypeptide chain following initiation. EF1 $\alpha$  is a GTP-binding protein. In EF1 $\alpha$ 's GTP-bound form, it brings an aminoacyl-tRNA to the ribosome's A site. The amino acid attached to the newly arrived aminoacyl-tRNA forms a peptide bond with the initiator methionine. 20 The GTP on EF1 $\alpha$  is hydrolyzed to GDP, and EF1 $\alpha$ -GDP dissociates from the ribosome. EF1 $\beta$   $\gamma$  binds EF1 $\alpha$ -GDP and induces the dissociation of GDP from EF1 $\alpha$ , allowing EF1 $\alpha$  to bind GTP and a new cycle to begin.

As subsequent aminoacyl-tRNAs are brought to the ribosome, EF-G, another GTP-binding protein, catalyzes the translocation of tRNAs from the A site to the P site and finally to the E site of the ribosome. This allows the processivity of translation.

#### Translation Termination

The release factor eRF carries out termination of translation. eRF recognizes stop codons in the mRNA, leading to the release of the polypeptide chain from the ribosome.

#### 30 Post-Translational Pathways

Proteins may be modified after translation by the addition of phosphate, sugar, prenyl, fatty acid, and other chemical groups. These modifications are often required for proper protein activity. Enzymes involved in post-translational modification include kinases, phosphatases, glycosyltransferases, and prenyltransferases. The conformation of proteins may also be modified after translation by the introduction and rearrangement of disulfide bonds (rearrangement catalyzed 35 by protein disulfide isomerase), the isomerization of proline sidechains by prolyl isomerase, and by

interactions with molecular chaperone proteins.

Proteins may also be cleaved by proteases. Such cleavage may result in activation, inactivation, or complete degradation of the protein. Proteases include serine proteases, cysteine proteases, aspartic proteases, and metalloproteases. Signal peptidase in the endoplasmic reticulum (ER) lumen cleaves the signal peptide from membrane or secretory proteins that are imported into the ER. Ubiquitin proteases are associated with the ubiquitin conjugation system (UCS), a major pathway for the degradation of cellular proteins in eukaryotic cells and some bacteria. The UCS mediates the elimination of abnormal proteins and regulates the half-lives of important regulatory proteins that control cellular processes such as gene transcription and cell cycle progression. In the UCS pathway, proteins targeted for degradation are conjugated to a ubiquitin, a small heat stable protein. Proteins involved in the UCS include ubiquitin-activating enzyme, ubiquitin-conjugating enzymes, ubiquitin-ligases, and ubiquitin C-terminal hydrolases. The ubiquitinated protein is then recognized and degraded by the proteasome, a large, multisubunit proteolytic enzyme complex, and ubiquitin is released for reutilization by ubiquitin protease.

#### 15 Lipid Metabolism

Lipids are water-insoluble, oily or greasy substances that are soluble in nonpolar solvents such as chloroform or ether. Neutral fats (triacylglycerols) serve as major fuels and energy stores. Polar lipids, such as phospholipids, sphingolipids, glycolipids, and cholesterol, are key structural components of cell membranes.

Lipid metabolism is involved in human diseases and disorders. In the arterial disease atherosclerosis, fatty lesions form on the inside of the arterial wall. These lesions promote the loss of arterial flexibility and the formation of blood clots (Guyton, A.C. Textbook of Medical Physiology (1991) W.B. Saunders Company, Philadelphia PA, pp.760-763). In Tay-Sachs disease, the GM<sub>2</sub> ganglioside (a sphingolipid) accumulates in lysosomes of the central nervous system due to a lack of the enzyme N-acetylhexosaminidase. Patients suffer nervous system degeneration leading to early death (Fauci, A.S. et al. (1998) Harrison's Principles of Internal Medicine McGraw-Hill, New York NY, p. 2171). The Niemann-Pick diseases are caused by defects in lipid metabolism. Niemann-Pick diseases types A and B are caused by accumulation of sphingomyelin (a sphingolipid) and other lipids in the central nervous system due to a defect in the enzyme sphingomyelinase, leading to neurodegeneration and lung disease. Niemann-Pick disease type C results from a defect in cholesterol transport, leading to the accumulation of sphingomyelin and cholesterol in lysosomes and a secondary reduction in sphingomyelinase activity. Neurological symptoms such as grand mal seizures, ataxia, and loss of previously learned speech, manifest 1-2 years after birth. A mutation in the NPC protein, which contains a putative cholesterol-sensing domain, was found in a mouse model of Niemann-Pick disease type C (Fauci, supra, p. 2175; Loftus, S.K. et al. (1997) *Science* 277:232-235). (Lipid metabolism is reviewed in Stryer, L. (1995) Biochemistry, W.H. Freeman and Company,

New York NY; Lehninger, A. (1982) Principles of Biochemistry Worth Publishers, Inc., New York NY; and ExPASy "Biochemical Pathways" index of Boehringer Mannheim World Wide Web site.)

### Fatty Acid Synthesis

Fatty acids are long-chain organic acids with a single carboxyl group and a long non-polar hydrocarbon tail. Long-chain fatty acids are essential components of glycolipids, phospholipids, and cholesterol, which are building blocks for biological membranes, and of triglycerides, which are biological fuel molecules. Long-chain fatty acids are also substrates for eicosanoid production, and are important in the functional modification of certain complex carbohydrates and proteins. 16-carbon and 18-carbon fatty acids are the most common.

Fatty acid synthesis occurs in the cytoplasm. In the first step, acetyl-Coenzyme A (CoA) carboxylase (ACC) synthesizes malonyl-CoA from acetyl-CoA and bicarbonate. The enzymes which catalyze the remaining reactions are covalently linked into a single polypeptide chain, referred to as the multifunctional enzyme fatty acid synthase (FAS). FAS catalyzes the synthesis of palmitate from acetyl-CoA and malonyl-CoA. FAS contains acetyl transferase, malonyl transferase,  $\beta$ -ketoacetyl synthase, acyl carrier protein,  $\beta$ -ketoacyl reductase, dehydratase, enoyl reductase, and thioesterase activities. The final product of the FAS reaction is the 16-carbon fatty acid palmitate. Further elongation, as well as unsaturation, of palmitate by accessory enzymes of the ER produces the variety of long chain fatty acids required by the individual cell. These enzymes include a NADH-cytochrome  $b_5$  reductase, cytochrome  $b_5$ , and a desaturase.

### Phospholipid and Triacylglycerol Synthesis

Triacylglycerols, also known as triglycerides and neutral fats, are major energy stores in animals. Triacylglycerols are esters of glycerol with three fatty acid chains. Glycerol-3-phosphate is produced from dihydroxyacetone phosphate by the enzyme glycerol phosphate dehydrogenase or from glycerol by glycerol kinase. Fatty acid-CoA's are produced from fatty acids by fatty acyl-CoA synthetases. Glycerol-3-phosphate is acylated with two fatty acyl-CoA's by the enzyme glycerol phosphate acyltransferase to give phosphatidate. Phosphatidate phosphatase converts phosphatidate to diacylglycerol, which is subsequently acylated to a triacylglycerol by the enzyme diglyceride acyltransferase. Phosphatidate phosphatase and diglyceride acyltransferase form a triacylglycerol synthetase complex bound to the ER membrane.

A major class of phospholipids are the phosphoglycerides, which are composed of a glycerol backbone, two fatty acid chains, and a phosphorylated alcohol. Phosphoglycerides are components of cell membranes. Principal phosphoglycerides are phosphatidyl choline, phosphatidyl ethanolamine, phosphatidyl serine, phosphatidyl inositol, and diphosphatidyl glycerol. Many enzymes involved in phosphoglyceride synthesis are associated with membranes (Meyers, R.A. (1995) Molecular Biology and Biotechnology, VCH Publishers Inc., New York NY, pp. 494-501). Phosphatidate is converted to CDP-diacylglycerol by the enzyme phosphatidate cytidyltransferase (ExPASy ENZYME EC

2.7.7.41). Transfer of the diacylglycerol group from CDP-diacylglycerol to serine to yield phosphatidyl serine, or to inositol to yield phosphatidyl inositol, is catalyzed by the enzymes CDP-diacylglycerol-serine O-phosphatidyltransferase and CDP-diacylglycerol-inositol 3-phosphatidyltransferase, respectively (ExPASy ENZYME EC 2.7.8.8; ExPASy ENZYME EC

- 5 2.7.8.11). The enzyme phosphatidyl serine decarboxylase catalyzes the conversion of phosphatidyl serine to phosphatidyl ethanolamine, using a pyruvate cofactor (Voelker, D.R. (1997) Biochim. Biophys. Acta 1348:236-244). Phosphatidyl choline is formed using diet-derived choline by the reaction of CDP-choline with 1,2-diacylglycerol, catalyzed by diacylglycerol cholinephosphotransferase (ExPASy ENZYME 2.7.8.2).

#### 10 Sterol, Steroid, and Isoprenoid Metabolism

- Cholesterol, composed of four fused hydrocarbon rings with an alcohol at one end, moderates the fluidity of membranes in which it is incorporated. In addition, cholesterol is used in the synthesis of steroid hormones such as cortisol, progesterone, estrogen, and testosterone. Bile salts derived from cholesterol facilitate the digestion of lipids. Cholesterol in the skin forms a barrier that prevents
- 15 excess water evaporation from the body. Farnesyl and geranylgeranyl groups, which are derived from cholesterol biosynthesis intermediates, are post-translationally added to signal transduction proteins such as ras and protein-targeting proteins such as rab. These modifications are important for the activities of these proteins (Guyton, supra; Stryer, supra, pp. 279-280, 691-702, 934).

- Mammals obtain cholesterol derived from both de novo biosynthesis and the diet. The liver
- 20 is the major site of cholesterol biosynthesis in mammals. Two acetyl-CoA molecules initially condense to form acetoacetyl-CoA, catalyzed by a thiolase. Acetoacetyl-CoA condenses with a third acetyl-CoA to form hydroxymethylglutaryl-CoA (HMG-CoA), catalyzed by HMG-CoA synthase. Conversion of HMG-CoA to cholesterol is accomplished via a series of enzymatic steps known as the mevalonate pathway. The rate-limiting step is the conversion of HMG-CoA to mevalonate by HMG-
- 25 CoA reductase. The drug lovastatin, a potent inhibitor of HMG-CoA reductase, is given to patients to reduce their serum cholesterol levels. Other mevalonate pathway enzymes include mevalonate kinase, phosphomevalonate kinase, diphosphomevalonate decarboxylase, isopentenyl diphosphate isomerase, dimethylallyl transferase, geranyl transferase, farnesyl-diphosphate farnesyltransferase, squalene monooxygenase, lanosterol synthase, lanosterol oxidase, and 7-dehydrocholesterol
- 30 reductase.

- Cholesterol is used in the synthesis of steroid hormones such as cortisol, progesterone, aldosterone, estrogen, and testosterone. First, cholesterol is converted to pregnenolone by cholesterol monooxygenases. The other steroid hormones are synthesized from pregnenolone by a series of enzyme-catalyzed reactions including oxidations, isomerizations, hydroxylations, reductions,
- 35 and demethylations. Examples of these enzymes include steroid  $\Delta$ -isomerase,  $3\beta$ -hydroxy- $\Delta^5$ -steroid dehydrogenase, steroid 21-monooxygenase, steroid 19-hydroxylase, and  $3\beta$ -hydroxysteroid

dehydrogenase. Cholesterol is also the precursor to vitamin D.

Numerous compounds contain 5-carbon isoprene units derived from the mevalonate pathway intermediate isopentenyl pyrophosphate. Isoprenoid groups are found in vitamin K, ubiquinone, retinal, dolichol phosphate (a carrier of oligosaccharides needed for N-linked glycosylation), and farnesyl and geranylgeranyl groups that modify proteins. Enzymes involved include farnesyl transferase, polyprenyl transferases, dolichyl phosphatase, and dolichyl kinase.

#### Sphingolipid Metabolism

Sphingolipids are an important class of membrane lipids that contain sphingosine, a long chain amino alcohol. They are composed of one long-chain fatty acid, one polar head alcohol, and sphingosine or sphingosine derivative. The three classes of sphingolipids are sphingomyelins, cerebroside, and gangliosides. Sphingomyelins, which contain phosphocholine or phosphoethanolamine as their head group, are abundant in the myelin sheath surrounding nerve cells. Galactocerebroside, which contains a glucose or galactose head group, are characteristic of the brain. Other cerebroside are found in nonneural tissues. Gangliosides, whose head groups contain multiple sugar units, are abundant in the brain, but are also found in nonneural tissues.

Sphingolipids are built on a sphingosine backbone. Sphingosine is acylated to ceramide by the enzyme sphingosine acetyltransferase. Ceramide and phosphatidyl choline are converted to sphingomyelin by the enzyme ceramide choline phosphotransferase. Cerebroside are synthesized by the linkage of glucose or galactose to ceramide by a transferase. Sequential addition of sugar residues to ceramide by transferase enzymes yields gangliosides.

#### Eicosanoid Metabolism

Eicosanoids, including prostaglandins, prostacyclin, thromboxanes, and leukotrienes, are 20-carbon molecules derived from fatty acids. Eicosanoids are signaling molecules which have roles in pain, fever, and inflammation. The precursor of all eicosanoids is arachidonate, which is generated from phospholipids by phospholipase A<sub>2</sub> and from diacylglycerols by diacylglycerol lipase. Leukotrienes are produced from arachidonate by the action of lipoxygenases. Prostaglandin synthase, reductases, and isomerases are responsible for the synthesis of the prostaglandins. Prostaglandins have roles in inflammation, blood flow, ion transport, synaptic transmission, and sleep. Prostacyclin and the thromboxanes are derived from a precursor prostaglandin by the action of prostacyclin synthase and thromboxane synthases, respectively.

#### Ketone Body Metabolism

Pairs of acetyl-CoA molecules derived from fatty acid oxidation in the liver can condense to form acetoacetyl-CoA, which subsequently forms acetoacetate, D-3-hydroxybutyrate, and acetone. These three products are known as ketone bodies. Enzymes involved in ketone body metabolism include HMG-CoA synthetase, HMG-CoA cleavage enzyme, D-3-hydroxybutyrate dehydrogenase, acetoacetate decarboxylase, and 3-ketoacyl-CoA transferase. Ketone bodies are a normal fuel supply

of the heart and renal cortex. Acetoacetate produced by the liver is transported to cells where the acetoacetate is converted back to acetyl-CoA and enters the citric acid cycle. In times of starvation, ketone bodies produced from stored triacylglycerols become an important fuel source, especially for the brain. Abnormally high levels of ketone bodies are observed in diabetics. Diabetic coma can  
5 result if ketone body levels become too great.

#### Lipid Mobilization

Within cells, fatty acids are transported by cytoplasmic fatty acid binding proteins (Online Mendelian Inheritance in Man (OMIM) \*134650 Fatty Acid-Binding Protein 1, Liver; FABP1). Diazepam binding inhibitor (DBI), also known as endozepine and acyl CoA-binding protein, is an  
10 endogenous  $\gamma$ -aminobutyric acid (GABA) receptor ligand which is thought to down-regulate the effects of GABA. DBI binds medium- and long-chain acyl-CoA esters with very high affinity and may function as an intracellular carrier of acyl-CoA esters (OMIM \*125950 Diazepam Binding Inhibitor; DBI; PROSITE PDOC00686 Acyl-CoA-binding protein signature).

Fat stored in liver and adipose triglycerides may be released by hydrolysis and transported in  
15 the blood. Free fatty acids are transported in the blood by albumin. Triacylglycerols and cholesterol esters in the blood are transported in lipoprotein particles. The particles consist of a core of hydrophobic lipids surrounded by a shell of polar lipids and apolipoproteins. The protein components serve in the solubilization of hydrophobic lipids and also contain cell-targeting signals. Lipoproteins include chylomicrons, chylomicron remnants, very-low-density lipoproteins (VLDL), intermediate-  
20 density lipoproteins (IDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL). There is a strong inverse correlation between the levels of plasma HDL and risk of premature coronary heart disease.

Triacylglycerols in chylomicrons and VLDL are hydrolyzed by lipoprotein lipases that line blood vessels in muscle and other tissues that use fatty acids. Cell surface LDL receptors bind LDL  
25 particles which are then internalized by endocytosis. Absence of the LDL receptor, the cause of the disease familial hypercholesterolemia, leads to increased plasma cholesterol levels and ultimately to atherosclerosis. Plasma cholesteryl ester transfer protein mediates the transfer of cholesteryl esters from HDL to apolipoprotein B-containing lipoproteins. Cholesteryl ester transfer protein is important in the reverse cholesterol transport system and may play a role in atherosclerosis (Yamashita, S. et al.  
30 (1997) Curr. Opin. Lipidol. 8:101-110). Macrophage scavenger receptors, which bind and internalize modified lipoproteins, play a role in lipid transport and may contribute to atherosclerosis (Greaves, D.R. et al. (1998) Curr. Opin. Lipidol. 9:425-432).

Proteins involved in cholesterol uptake and biosynthesis are tightly regulated in response to cellular cholesterol levels. The sterol regulatory element binding protein (SREBP) is a sterol-  
35 responsive transcription factor. Under normal cholesterol conditions, SREBP resides in the ER membrane. When cholesterol levels are low, a regulated cleavage of SREBP occurs which releases

the extracellular domain of the protein. This cleaved domain is then transported to the nucleus where it activates the transcription of the LDL receptor gene, and genes encoding enzymes of cholesterol synthesis, by binding the sterol regulatory element (SRE) upstream of the genes (Yang, J. et al. (1995) *J. Biol. Chem.* 270:12152-12161). Regulation of cholesterol uptake and biosynthesis also occurs via the oxysterol-binding protein (OSBP). OSBP is a high-affinity intracellular receptor for a variety of oxysterols that down-regulate cholesterol synthesis and stimulate cholesterol esterification (Lagace, T.A. et al. (1997) *Biochem. J.* 326:205-213).

#### Beta-oxidation

Mitochondrial and peroxisomal beta-oxidation enzymes degrade saturated and unsaturated fatty acids by sequential removal of two-carbon units from CoA-activated fatty acids. The main beta-oxidation pathway degrades both saturated and unsaturated fatty acids while the auxiliary pathway performs additional steps required for the degradation of unsaturated fatty acids.

The pathways of mitochondrial and peroxisomal beta-oxidation use similar enzymes, but have different substrate specificities and functions. Mitochondria oxidize short-, medium-, and long-chain fatty acids to produce energy for cells. Mitochondrial beta-oxidation is a major energy source for cardiac and skeletal muscle. In liver, it provides ketone bodies to the peripheral circulation when glucose levels are low as in starvation, endurance exercise, and diabetes (Eaton, S. et al. (1996) *Biochem. J.* 320:345-357). Peroxisomes oxidize medium-, long-, and very-long-chain fatty acids, dicarboxylic fatty acids, branched fatty acids, prostaglandins, xenobiotics, and bile acid intermediates. The chief roles of peroxisomal beta-oxidation are to shorten toxic lipophilic carboxylic acids to facilitate their excretion and to shorten very-long-chain fatty acids prior to mitochondrial beta-oxidation (Mannaerts, G.P. and P.P. van Veldhoven (1993) *Biochimie* 75:147-158).

Enzymes involved in beta-oxidation include acyl CoA synthetase, carnitine acyltransferase, acyl CoA dehydrogenases, enoyl CoA hydratases, L-3-hydroxyacyl CoA dehydrogenase,  $\beta$ -ketothiolase, 2,4-dienoyl CoA reductase, and isomerase.

#### Lipid Cleavage and Degradation

Triglycerides are hydrolyzed to fatty acids and glycerol by lipases. Lysophospholipases (LPLs) are widely distributed enzymes that metabolize intracellular lipids, and occur in numerous isoforms. Small isoforms, approximately 15-30 kD, function as hydrolases; large isoforms, those exceeding 60 kD, function both as hydrolases and transacylases. A particular substrate for LPLs, lysophosphatidylcholine, causes lysis of cell membranes when it is formed or imported into a cell. LPLs are regulated by lipid factors including acylcarnitine, arachidonic acid, and phosphatidic acid. These lipid factors are signaling molecules important in numerous pathways, including the inflammatory response. (Anderson, R. et al. (1994) *Toxicol. Appl. Pharmacol.* 125:176-183; Selle, H. et al. (1993); *Eur. J. Biochem.* 212:411-416.)

The secretory phospholipase A<sub>2</sub> (PLA<sub>2</sub>) superfamily comprises a number of heterogeneous enzymes whose common feature is to hydrolyze the sn-2 fatty acid acyl ester bond of phosphoglycerides. Hydrolysis of the glycerophospholipids releases free fatty acids and lysophospholipids. PLA<sub>2</sub> activity generates precursors for the biosynthesis of biologically active lipids, hydroxy fatty acids, and platelet-activating factor. PLA<sub>2</sub> hydrolysis of the sn-2 ester bond in phospholipids generates free fatty acids, such as arachidonic acid and lysophospholipids.

#### Carbon and Carbohydrate Metabolism

Carbohydrates, including sugars or saccharides, starch, and cellulose, are aldehyde or ketone compounds with multiple hydroxyl groups. The importance of carbohydrate metabolism is demonstrated by the sensitive regulatory system in place for maintenance of blood glucose levels. Two pancreatic hormones, insulin and glucagon, promote increased glucose uptake and storage by cells, and increased glucose release from cells, respectively. Carbohydrates have three important roles in mammalian cells. First, carbohydrates are used as energy stores, fuels, and metabolic intermediates. Carbohydrates are broken down to form energy in glycolysis and are stored as glycogen for later use. Second, the sugars deoxyribose and ribose form part of the structural support of DNA and RNA, respectively. Third, carbohydrate modifications are added to secreted and membrane proteins and lipids as they traverse the secretory pathway. Cell surface carbohydrate-containing macromolecules, including glycoproteins, glycolipids, and transmembrane proteoglycans, mediate adhesion with other cells and with components of the extracellular matrix. The extracellular matrix is comprised of diverse glycoproteins, glycosaminoglycans (GAGs), and carbohydrate-binding proteins which are secreted from the cell and assembled into an organized meshwork in close association with the cell surface. The interaction of the cell with the surrounding matrix profoundly influences cell shape, strength, flexibility, motility, and adhesion. These dynamic properties are intimately associated with signal transduction pathways controlling cell proliferation and differentiation, tissue construction, and embryonic development.

Carbohydrate metabolism is altered in several disorders including diabetes mellitus, hyperglycemia, hypoglycemia, galactosemia, galactokinase deficiency, and UDP-galactose-4-epimerase deficiency (Fauci, A.S. et al. (1998) Harrison's Principles of Internal Medicine, McGraw-Hill, New York NY, pp. 2208-2209). Altered carbohydrate metabolism is associated with cancer. Reduced GAG and proteoglycan expression is associated with human lung carcinomas (Nackaerts, K. et al. (1997) *Int. J. Cancer* 74:335-345). The carbohydrate determinants sialyl Lewis A and sialyl Lewis X are frequently expressed on human cancer cells (Kannagi, R. (1997) *Glycoconj. J.* 14:577-584). Alterations of the N-linked carbohydrate core structure of cell surface glycoproteins are linked to colon and pancreatic cancers (Schwarz, R.E. et al. (1996) *Cancer Lett.* 107:285-291). Reduced expression of the Sda blood group carbohydrate structure in cell surface glycolipids and glycoproteins is observed in gastrointestinal cancer (Dohi, T. et al. (1996) *Int. J. Cancer* 67:626-663). (Carbon and



carbohydrate metabolism is reviewed in Stryer, L. (1995) Biochemistry W.H. Freeman and Company, New York NY; Lehninger, A.L. (1982) Principles of Biochemistry Worth Publishers Inc., New York NY; and Lodish, H. et al. (1995) Molecular Cell Biology Scientific American Books, New York NY.)

#### Glycolysis

- 5 Enzymes of the glycolytic pathway convert the sugar glucose to pyruvate while simultaneously producing ATP. The pathway also provides building blocks for the synthesis of cellular components such as long-chain fatty acids. After glycolysis, pyruvate is converted to acetyl-Coenzyme A, which, in aerobic organisms, enters the citric acid cycle. Glycolytic enzymes include hexokinase, phosphoglucose isomerase, phosphofructokinase, aldolase, triose phosphate isomerase, 10 glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglyceromutase, enolase, and pyruvate kinase. Of these, phosphofructokinase, hexokinase, and pyruvate kinase are important in regulating the rate of glycolysis.

#### Gluconeogenesis

- Gluconeogenesis is the synthesis of glucose from noncarbohydrate precursors such as lactate 15 and amino acids. The pathway, which functions mainly in times of starvation and intense exercise, occurs mostly in the liver and kidney. Responsible enzymes include pyruvate carboxylase, phosphoenolpyruvate carboxykinase, fructose 1,6-bisphosphatase, and glucose-6-phosphatase.

#### Pentose Phosphate Pathway

- Pentose phosphate pathway enzymes are responsible for generating the reducing agent 20 NADPH, while at the same time oxidizing glucose-6-phosphate to ribose-5-phosphate. Ribose-5-phosphate and its derivatives become part of important biological molecules such as ATP, Coenzyme A, NAD<sup>+</sup>, FAD, RNA, and DNA. The pentose phosphate pathway has both oxidative and non-oxidative branches. The oxidative branch steps, which are catalyzed by the enzymes glucose-6-phosphate dehydrogenase, lactonase, and 6-phosphogluconate dehydrogenase, convert glucose-6-phosphate and NADP<sup>+</sup> to ribulose-6-phosphate and NADPH. 25 The non-oxidative branch steps, which are catalyzed by the enzymes phosphopentose isomerase, phosphopentose epimerase, transketolase, and transaldolase, allow the interconversion of three-, four-, five-, six-, and seven-carbon sugars.

#### Glucuronate Metabolism

- Glucuronate is a monosaccharide which, in the form of D-glucuronic acid, is found in the 30 GAGs chondroitin and dermatan. D-glucuronic acid is also important in the detoxification and excretion of foreign organic compounds such as phenol. Enzymes involved in glucuronate metabolism include UDP-glucose dehydrogenase and glucuronate reductase.

#### Disaccharide Metabolism

- Disaccharides must be hydrolyzed to monosaccharides to be digested. Lactose, a 35 disaccharide found in milk, is hydrolyzed to galactose and glucose by the enzyme lactase. Maltose is derived from plant starch and is hydrolyzed to glucose by the enzyme maltase. Sucrose is derived

from plants and is hydrolyzed to glucose and fructose by the enzyme sucrase. Trehalose, a disaccharide found mainly in insects and mushrooms, is hydrolyzed to glucose by the enzyme trehalase (OMIM \*275360 Trehalase; Ruf, J. et al. (1990) J. Biol. Chem. 265:15034-15039). Lactase, maltase, sucrase, and trehalase are bound to mucosal cells lining the small intestine, where they participate in the digestion of dietary disaccharides. The enzyme lactose synthetase, composed of the catalytic subunit galactosyltransferase and the modifier subunit  $\alpha$ -lactalbumin, converts UDP-galactose and glucose to lactose in the mammary glands.

#### Glycogen, Starch, and Chitin Metabolism

Glycogen is the storage form of carbohydrates in mammals. Mobilization of glycogen maintains glucose levels between meals and during muscular activity. Glycogen is stored mainly in the liver and in skeletal muscle in the form of cytoplasmic granules. These granules contain enzymes that catalyze the synthesis and degradation of glycogen, as well as enzymes that regulate these processes. Enzymes that catalyze the degradation of glycogen include glycogen phosphorylase, a transferase,  $\alpha$ -1,6-glucosidase, and phosphoglucomutase. Enzymes that catalyze the synthesis of glycogen include UDP-glucose pyrophosphorylase, glycogen synthetase, a branching enzyme, and nucleoside diphosphokinase. The enzymes of glycogen synthesis and degradation are tightly regulated by the hormones insulin, glucagon, and epinephrine. Starch, a plant-derived polysaccharide, is hydrolyzed to maltose, maltotriose, and  $\alpha$ -dextrin by  $\alpha$ -amylase, an enzyme secreted by the salivary glands and pancreas. Chitin is a polysaccharide found in insects and crustacea. A chitotriosidase is secreted by macrophages and may play a role in the degradation of chitin-containing pathogens (Boot, R.G. et al. (1995) J. Biol. Chem. 270:26252-26256).

#### Peptidoglycans and Glycosaminoglycans

Glycosaminoglycans (GAGs) are anionic linear unbranched polysaccharides composed of repetitive disaccharide units. These repetitive units contain a derivative of an amino sugar, either glucosamine or galactosamine. GAGs exist free or as part of proteoglycans, large molecules composed of a core protein attached to one or more GAGs. GAGs are found on the cell surface, inside cells, and in the extracellular matrix. Changes in GAG levels are associated with several autoimmune diseases including autoimmune thyroid disease, autoimmune diabetes mellitus, and systemic lupus erythematosus (Hansen, C. et al. (1996) Clin. Exp. Rheum. 14 (Suppl. 15):S59-S67). GAGs include chondroitin sulfate, keratan sulfate, heparin, heparan sulfate, dermatan sulfate, and hyaluronan.

The GAG hyaluronan (HA) is found in the extracellular matrix of many cells, especially in soft connective tissues, and is abundant in synovial fluid (Pitsillides, A.A. et al. (1993) Int. J. Exp. Pathol. 74:27-34). HA seems to play important roles in cell regulation, development, and differentiation (Laurent, T.C. and J.R. Fraser (1992) FASEB J. 6:2397-2404). Hyaluronidase is an enzyme that degrades HA to oligosaccharides. Hyaluronidases may function in cell adhesion,

infection, angiogenesis, signal transduction, reproduction, cancer, and inflammation.

Proteoglycans, also known as peptidoglycans, are found in the extracellular matrix of connective tissues such as cartilage and are essential for distributing the load in weight-bearing joints. Cell-surface-attached proteoglycans anchor cells to the extracellular matrix. Both extracellular and cell-surface proteoglycans bind growth factors, facilitating their binding to cell-surface receptors and subsequent triggering of signal transduction pathways.

#### Amino Acid and Nitrogen Metabolism

$\text{NH}_4^+$  is assimilated into amino acids by the actions of two enzymes, glutamate dehydrogenase and glutamine synthetase. The carbon skeletons of amino acids come from the intermediates of glycolysis, the pentose phosphate pathway, or the citric acid cycle. Of the twenty amino acids used in proteins, humans can synthesize only thirteen (nonessential amino acids). The remaining nine must come from the diet (essential amino acids). Enzymes involved in nonessential amino acid biosynthesis include glutamate kinase dehydrogenase, pyrroline carboxylate reductase, asparagine synthetase, phenylalanine oxygenase, methionine adenosyltransferase, adenosylhomocysteinase, cystathionine  $\beta$ -synthase, cystathionine  $\gamma$ -lyase, phosphoglycerate dehydrogenase, phosphoserine transaminase, phosphoserine phosphatase, serine hydroxymethyltransferase, and glycine synthase.

Metabolism of amino acids takes place almost entirely in the liver, where the amino group is removed by aminotransferases (transaminases), for example, alanine aminotransferase. The amino group is transferred to  $\alpha$ -ketoglutarate to form glutamate. Glutamate dehydrogenase converts glutamate to  $\text{NH}_4^+$  and  $\alpha$ -ketoglutarate.  $\text{NH}_4^+$  is converted to urea by the urea cycle which is catalyzed by the enzymes arginase, ornithine transcarbamoylase, arginosuccinate synthetase, and arginosuccinase. Carbamoyl phosphate synthetase is also involved in urea formation. Enzymes involved in the metabolism of the carbon skeleton of amino acids include serine dehydratase, asparaginase, glutaminase, propionyl CoA carboxylase, methylmalonyl CoA mutase, branched-chain  $\alpha$ -keto dehydrogenase complex, isovaleryl CoA dehydrogenase,  $\beta$ -methylcrotonyl CoA carboxylase, phenylalanine hydroxylase, p-hydroxyphenylpyruvate hydroxylase, and homogentisate oxidase.

Polyamines, which include spermidine, putrescine, and spermine, bind tightly to nucleic acids and are abundant in rapidly proliferating cells. Enzymes involved in polyamine synthesis include ornithine decarboxylase.

Diseases involved in amino acid and nitrogen metabolism include hyperammonemia, carbamoyl phosphate synthetase deficiency, urea cycle enzyme deficiencies, methylmalonic aciduria, maple syrup disease, alcaptonuria, and phenylketonuria.

#### Energy Metabolism

Cells derive energy from metabolism of ingested compounds that may be roughly categorized as carbohydrates, fats, or proteins. Energy is also stored in polymers such as triglycerides (fats) and

glycogen (carbohydrates). Metabolism proceeds along separate reaction pathways connected by key intermediates such as acetyl coenzyme A (acetyl-CoA). Metabolic pathways feature anaerobic and aerobic degradation, coupled with the energy-requiring reactions such as phosphorylation of adenosine diphosphate (ADP) to the triphosphate (ATP) or analogous phosphorylations of guanosine (GDP/GTP), uridine (UDP/UTP), or cytidine (CDP/CTP). Subsequent dephosphorylation of the triphosphate drives reactions needed for cell maintenance, growth, and proliferation.

Digestive enzymes convert carbohydrates and sugars to glucose; fructose and galactose are converted in the liver to glucose. Enzymes involved in these conversions include galactose-1-phosphate uridyl transferase and UDP-galactose-4 epimerase. In the cytoplasm, glycolysis converts glucose to pyruvate in a series of reactions coupled to ATP synthesis.

Pyruvate is transported into the mitochondria and converted to acetyl-CoA for oxidation via the citric acid cycle, involving pyruvate dehydrogenase components, dihydrolipoyl transacetylase, and dihydrolipoyl dehydrogenase. Enzymes involved in the citric acid cycle include: citrate synthetase, aconitases, isocitrate dehydrogenase, alpha-ketoglutarate dehydrogenase complex including transsuccinylases, succinyl CoA synthetase, succinate dehydrogenase, fumarases, and malate dehydrogenase. Acetyl CoA is oxidized to CO<sub>2</sub> with concomitant formation of NADH, FADH<sub>2</sub>, and GTP. In oxidative phosphorylation, the transport of electrons from NADH and FADH<sub>2</sub> to oxygen by dehydrogenases is coupled to the synthesis of ATP from ADP and P<sub>i</sub> by the F<sub>0</sub>F<sub>1</sub> ATPase complex in the mitochondrial inner membrane. Enzyme complexes responsible for electron transport and ATP synthesis include the F<sub>0</sub>F<sub>1</sub> ATPase complex, ubiquinone(CoQ)-cytochrome c reductase, ubiquinone reductase, cytochrome b, cytochrome c<sub>1</sub>, FeS protein, and cytochrome c oxidase.

Triglycerides are hydrolyzed to fatty acids and glycerol by lipases. Glycerol is then phosphorylated to glycerol-3-phosphate by glycerol kinase and glycerol phosphate dehydrogenase, and degraded by the glycolysis. Fatty acids are transported into the mitochondria as fatty acyl-carnitine esters and undergo oxidative degradation.

In addition to metabolic disorders such as diabetes and obesity, disorders of energy metabolism are associated with cancers (Dorward, A. et al. (1997) J. Bioenerg. Biomembr. 29:385-392), autism (Lombard, J. (1998) Med. Hypotheses 50:497-500), neurodegenerative disorders (Alexi, T. et al. (1998) Neuroreport 9:R57-64), and neuromuscular disorders (DiMauro, S. et al. (1998) Biochim. Biophys. Acta 1366:199-210). The myocardium is heavily dependent on oxidative metabolism, so metabolic dysfunction often leads to heart disease (DiMauro, S. and M. Hirano (1998) Curr. Opin. Cardiol. 13:190-197).

For a review of energy metabolism enzymes and intermediates, see Stryer, L. et al. (1995) Biochemistry, W.H. Freeman and Co., San Francisco CA, pp. 443-652. For a review of energy metabolism regulation, see Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American

Books, New York NY, pp. 744-770.

### Cofactor Metabolism

Cofactors, including coenzymes and prosthetic groups, are small molecular weight inorganic or organic compounds that are required for the action of an enzyme. Many cofactors contain  
5 vitamins as a component. Cofactors include thiamine pyrophosphate, flavin adenine dinucleotide, flavin mononucleotide, nicotinamide adenine dinucleotide, pyridoxal phosphate, coenzyme A, tetrahydrofolate, lipoamide, and heme. The vitamins biotin and cobalamin are associated with enzymes as well. Heme, a prosthetic group found in myoglobin and hemoglobin, consists of  
10 protoporphyrin group bound to iron. Porphyrin groups contain four substituted pyrroles covalently joined in a ring, often with a bound metal atom. Enzymes involved in porphyrin synthesis include  $\delta$ -aminolevulinate synthase,  $\delta$ -aminolevulinate dehydrase, porphobilinogen deaminase, and cosynthase. Deficiencies in heme formation cause porphyrias. Heme is broken down as a part of erythrocyte turnover. Enzymes involved in heme degradation include heme oxygenase and biliverdin reductase.

Iron is a required cofactor for many enzymes. Besides the heme-containing enzymes, iron is  
15 found in iron-sulfur clusters in proteins including aconitase, succinate dehydrogenase, and NADH-Q reductase. Iron is transported in the blood by the protein transferrin. Binding of transferrin to the transferrin receptor on cell surfaces allows uptake by receptor mediated endocytosis. Cytosolic iron is bound to ferritin protein.

A molybdenum-containing cofactor (molybdopterin) is found in enzymes including sulfite  
20 oxidase, xanthine dehydrogenase, and aldehyde oxidase. Molybdopterin biosynthesis is performed by two molybdenum cofactor synthesizing enzymes. Deficiencies in these enzymes cause mental retardation and lens dislocation. Other diseases caused by defects in cofactor metabolism include pernicious anemia and methylmalonic aciduria.

### Secretion and Trafficking

Eukaryotic cells are bound by a lipid bilayer membrane and subdivided into functionally  
25 distinct, membrane bound compartments. The membranes maintain the essential differences between the cytosol, the extracellular environment, and the lumenal space of each intracellular organelle. As lipid membranes are highly impermeable to most polar molecules, transport of essential nutrients, metabolic waste products, cell signaling molecules, macromolecules and proteins across lipid  
30 membranes and between organelles must be mediated by a variety of transport-associated molecules.

### Protein Trafficking

In eukaryotes, some proteins are synthesized on ER-bound ribosomes, co-translationally imported into the ER, delivered from the ER to the Golgi complex for post-translational processing and sorting, and transported from the Golgi to specific intracellular and extracellular destinations.  
35 All cells possess a constitutive transport process which maintains homeostasis between the cell and its environment. In many differentiated cell types, the basic machinery is modified to carry out

specific transport functions. For example, in endocrine glands, hormones and other secreted proteins are packaged into secretory granules for regulated exocytosis to the cell exterior. In macrophage, foreign extracellular material is engulfed (phagocytosis) and delivered to lysosomes for degradation. In fat and muscle cells, glucose transporters are stored in vesicles which fuse with the plasma membrane only in response to insulin stimulation.

#### The Secretory Pathway

Synthesis of most integral membrane proteins, secreted proteins, and proteins destined for the lumen of a particular organelle occurs on ER-bound ribosomes. These proteins are co-translationally imported into the ER. The proteins leave the ER via membrane-bound vesicles which bud off the ER at specific sites and fuse with each other (homotypic fusion) to form the ER-Golgi Intermediate Compartment (ERGIC). The ERGIC matures progressively through the *cis*, *medial*, and *trans* cisternal stacks of the Golgi, modifying the enzyme composition by retrograde transport of specific Golgi enzymes. In this way, proteins moving through the Golgi undergo post-translational modification, such as glycosylation. The final Golgi compartment is the Trans-Golgi Network (TGN), where both membrane and luminal proteins are sorted for their final destination. Transport vesicles destined for intracellular compartments, such as the lysosome, bud off the TGN. What remains is a secretory vesicle which contains proteins destined for the plasma membrane, such as receptors, adhesion molecules, and ion channels, and secretory proteins, such as hormones, neurotransmitters, and digestive enzymes. Secretory vesicles eventually fuse with the plasma membrane (Glick, B.S. and V. Malhotra (1998) Cell 95:883-889).

The secretory process can be constitutive or regulated. Most cells have a constitutive pathway for secretion, whereby vesicles derived from maturation of the TGN require no specific signal to fuse with the plasma membrane. In many cells, such as endocrine cells, digestive cells, and neurons, vesicle pools derived from the TGN collect in the cytoplasm and do not fuse with the plasma membrane until they are directed to by a specific signal.

#### Endocytosis

Endocytosis, wherein cells internalize material from the extracellular environment, is essential for transmission of neuronal, metabolic, and proliferative signals; uptake of many essential nutrients; and defense against invading organisms. Most cells exhibit two forms of endocytosis. The first, phagocytosis, is an actin-driven process exemplified in macrophage and neutrophils. Material to be endocytosed contacts numerous cell surface receptors which stimulate the plasma membrane to extend and surround the particle, enclosing it in a membrane-bound phagosome. In the mammalian immune system, IgG-coated particles bind Fc receptors on the surface of phagocytic leukocytes. Activation of the Fc receptors initiates a signal cascade involving src-family cytosolic kinases and the monomeric GTP-binding (G) protein Rho. The resulting actin reorganization leads to phagocytosis of the particle. This process is an important component of the humoral immune response, allowing the

processing and presentation of bacterial-derived peptides to antigen-specific T-lymphocytes.

The second form of endocytosis, pinocytosis, is a more generalized uptake of material from the external milieu. Like phagocytosis, pinocytosis is activated by ligand binding to cell surface receptors. Activation of individual receptors stimulates an internal response that includes

5 coalescence of the receptor-ligand complexes and formation of clathrin-coated pits. Invagination of the plasma membrane at clathrin-coated pits produces an endocytic vesicle within the cell cytoplasm. These vesicles undergo homotypic fusion to form an early endosomal (EE) compartment. The tubulovesicular EE serves as a sorting site for incoming material. ATP-driven proton pumps in the EE membrane lowers the pH of the EE lumen (pH 6.3-6.8). The acidic environment causes many

10 ligands to dissociate from their receptors. The receptors, along with membrane and other integral membrane proteins, are recycled back to the plasma membrane by budding off the tubular extensions of the EE in recycling vesicles (RV). This selective removal of recycled components produces a carrier vesicle containing ligand and other material from the external environment. The carrier vesicle fuses with TGN-derived vesicles which contain hydrolytic enzymes. The acidic environment

15 of the resulting late endosome (LE) activates the hydrolytic enzymes which degrade the ligands and other material. As digestion takes place, the LE fuses with the lysosome where digestion is completed (Mellman, I. (1996) *Annu. Rev. Cell Dev. Biol.* 12:575-625).

Recycling vesicles may return directly to the plasma membrane. Receptors internalized and returned directly to the plasma membrane have a turnover rate of 2-3 minutes. Some RVs undergo

20 microtubule-directed relocation to a perinuclear site, from which they then return to the plasma membrane. Receptors following this route have a turnover rate of 5-10 minutes. Still other RVs are retained within the cell until an appropriate signal is received (Mellman, *supra*; and James, D.E. et al. (1994) *Trends Cell Biol.* 4:120-126).

#### Vesicle Formation

Several steps in the transit of material along the secretory and endocytic pathways require the

25 formation of transport vesicles. Specifically, vesicles form at the transitional endoplasmic reticulum (tER), the rim of Golgi cisternae, the face of the Trans-Golgi Network (TGN), the plasma membrane (PM), and tubular extensions of the endosomes. The process begins with the budding of a vesicle out of the donor membrane. The membrane-bound vesicle contains proteins to be transported and is

30 surrounded by a protective coat made up of protein subunits recruited from the cytosol. The initial budding and coating processes are controlled by a cytosolic ras-like GTP-binding protein, ADP-ribosylating factor (Arf), and adapter proteins (AP). Different isoforms of both Arf and AP are involved at different sites of budding. Another small G-protein, dynamin, forms a ring complex around the neck of the forming vesicle and may provide the mechanochemical force to accomplish the

35 final step of the budding process. The coated vesicle complex is then transported through the cytosol. During the transport process, Arf-bound GTP is hydrolyzed to GDP and the coat dissociates from the

transport vesicle (West, M.A. et al. (1997) J. Cell Biol. 138:1239-1254). Two different classes of coat protein have also been identified. Clathrin coats form on the TGN and PM surfaces, whereas coatomer or COP coats form on the ER and Golgi. COP coats can further be distinguished as COPI, involved in retrograde traffic through the Golgi and from the Golgi to the ER, and COPII, involved in anterograde traffic from the ER to the Golgi (Mellman, supra). The COP coat consists of two major components, a G-protein (Arf or Sar) and coat protomer (coatomer). Coatomer is an equimolar complex of seven proteins, termed alpha-, beta-, beta'-, gamma-, delta-, epsilon- and zeta-COP. (Harter, C. and F.T. Wieland (1998) Proc. Natl. Acad. Sci. USA 95:11649-11654.)

#### Membrane Fusion

Transport vesicles undergo homotypic or heterotypic fusion in the secretory and endocytotic pathways. Molecules required for appropriate targeting and fusion of vesicles with their target membrane include proteins incorporated in the vesicle membrane, the target membrane, and proteins recruited from the cytosol. During budding of the vesicle from the donor compartment, an integral membrane protein, VAMP (vesicle-associated membrane protein) is incorporated into the vesicle. Soon after the vesicle uncoats, a cytosolic prenylated GTP-binding protein, Rab (a member of the Ras superfamily), is inserted into the vesicle membrane. GTP-bound Rab proteins are directed into nascent transport vesicles where they interact with VAMP. Following vesicle transport, GTPase activating proteins (GAPs) in the target membrane convert Rab proteins to the GDP-bound form. A cytosolic protein, guanine-nucleotide dissociation inhibitor (GDI) helps return GDP-bound Rab proteins to their membrane of origin. Several Rab isoforms have been identified and appear to associate with specific compartments within the cell. Rab proteins appear to play a role in mediating the function of a viral gene, Rev, which is essential for replication of HIV-1, the virus responsible for AIDS (Flavell, R.A. et al. (1996) Proc. Natl. Acad. Sci. USA 93:4421-4424).

Docking of the transport vesicle with the target membrane involves the formation of a complex between the vesicle SNAP receptor (v-SNARE), target membrane (t-) SNAREs, and certain other membrane and cytosolic proteins. Many of these other proteins have been identified although their exact functions in the docking complex remain uncertain (Tellam, J.T. et al. (1995) J. Biol. Chem. 270:5857-5863; and Hata, Y. and T.C. Sudhof (1995) J. Biol. Chem. 270:13022-13028). N-ethylmaleimide sensitive factor (NSF) and soluble NSF-attachment protein ( $\alpha$ -SNAP and  $\beta$ -SNAP) are two such proteins that are conserved from yeast to man and function in most intracellular membrane fusion reactions. Sec1 represents a family of yeast proteins that function at many different stages in the secretory pathway including membrane fusion. Recently, mammalian homologs of Sec1, called Munc-18 proteins, have been identified (Katagiri, H. et al. (1995) J. Biol. Chem. 270:4963-4966; Hata et al. supra).

The SNARE complex involves three SNARE molecules, one in the vesicular membrane and two in the target membrane. Synaptotagmin is an integral membrane protein in the synaptic vesicle



which associates with the t-SNARE syntaxin in the docking complex. Synaptotagmin binds calcium in a complex with negatively charged phospholipids, which allows the cytosolic SNAP protein to displace synaptotagmin from syntaxin and fusion to occur. Thus, synaptotagmin is a negative regulator of fusion in the neuron (Littleton, J.T. et al. (1993) Cell 74:1125-1134). The most abundant  
5 membrane protein of synaptic vesicles appears to be the glycoprotein synaptophysin, a 38 kDa protein with four transmembrane domains.

Specificity between a vesicle and its target is derived from the v-SNARE, t-SNAREs, and associated proteins involved. Different isoforms of SNAREs and Rabs show distinct cellular and subcellular distributions. VAMP-1/synaptobrevin, membrane-anchored synaptosome-associated  
10 protein of 25 kDa (SNAP-25), syntaxin-1, Rab3A, Rab15, and Rab23 are predominantly expressed in the brain and nervous system. Different syntaxin, VAMP, and Rab proteins are associated with distinct subcellular compartments and their vesicular carriers.

#### Nuclear Transport

Transport of proteins and RNA between the nucleus and the cytoplasm occurs through  
15 nuclear pore complexes (NPCs). NPC-mediated transport occurs in both directions through the nuclear envelope. All nuclear proteins are imported from the cytoplasm, their site of synthesis. tRNA and mRNA are exported from the nucleus, their site of synthesis, to the cytoplasm, their site of function. Processing of small nuclear RNAs involves export into the cytoplasm, assembly with proteins and modifications such as hypermethylation to produce small nuclear ribonuclear proteins  
20 (snRNPs), and subsequent import of the snRNPs back into the nucleus. The assembly of ribosomes requires the initial import of ribosomal proteins from the cytoplasm, their incorporation with RNA into ribosomal subunits, and export back to the cytoplasm. (Görllich, D. and I.W. Mattaj (1996) Science 271:1513-1518.)

The transport of proteins and mRNAs across the NPC is selective, dependent on nuclear  
25 localization signals, and generally requires association with nuclear transport factors. Nuclear localization signals (NLS) consist of short stretches of amino acids enriched in basic residues. NLS are found on proteins that are targeted to the nucleus, such as the glucocorticoid receptor. The NLS is recognized by the NLS receptor, importin, which then interacts with the monomeric GTP-binding protein Ran. This NLS protein/receptor/Ran complex navigates the nuclear pore with the help of the  
30 homodimeric protein nuclear transport factor 2 (NTF2). NTF2 binds the GDP-bound form of Ran and to multiple proteins of the nuclear pore complex containing FXFG repeat motifs, such as p62. (Paschal, B. et al. (1997) J. Biol. Chem. 272:21534-21539; and Wong, D.H. et al. (1997) Mol. Cell Biol. 17:3755-3767). Some proteins are dissociated before nuclear mRNAs are transported across the NPC while others are dissociated shortly after nuclear mRNA transport across the NPC and are  
35 reimported into the nucleus.

#### Disease Correlation

The etiology of numerous human diseases and disorders can be attributed to defects in the transport or secretion of proteins. For example, abnormal hormonal secretion is linked to disorders such as diabetes insipidus (vasopressin), hyper- and hypoglycemia (insulin, glucagon), Grave's disease and goiter (thyroid hormone), and Cushing's and Addison's diseases (adrenocorticotrophic hormone, ACTH). Moreover, cancer cells secrete excessive amounts of hormones or other biologically active peptides. Disorders related to excessive secretion of biologically active peptides by tumor cells include fasting hypoglycemia due to increased insulin secretion from insulinoma-islet cell tumors; hypertension due to increased epinephrine and norepinephrine secreted from pheochromocytomas of the adrenal medulla and sympathetic paraganglia; and carcinoid syndrome, which is characterized by abdominal cramps, diarrhea, and valvular heart disease caused by excessive amounts of vasoactive substances such as serotonin, bradykinin, histamine, prostaglandins, and polypeptide hormones, secreted from intestinal tumors. Biologically active peptides that are ectopically synthesized in and secreted from tumor cells include ACTH and vasopressin (lung and pancreatic cancers); parathyroid hormone (lung and bladder cancers); calcitonin (lung and breast cancers); and thyroid-stimulating hormone (medullary thyroid carcinoma). Such peptides may be useful as diagnostic markers for tumorigenesis (Schwartz, M.Z. (1997) *Semin. Pediatr. Surg.* 3:141-146; and Said, S.I. and G.R. Faloona (1975) *N. Engl. J. Med.* 293:155-160).

Defective nuclear transport may play a role in cancer. The BRCA1 protein contains three potential NLSs which interact with importin alpha, and is transported into the nucleus by the importin/NPC pathway. In breast cancer cells the BRCA1 protein is aberrantly localized in the cytoplasm. The mislocation of the BRCA1 protein in breast cancer cells may be due to a defect in the NPC nuclear import pathway (Chen, C.F. et al. (1996) *J. Biol. Chem.* 271:32863-32868).

It has been suggested that in some breast cancers, the tumor-suppressing activity of p53 is inactivated by the sequestration of the protein in the cytoplasm, away from its site of action in the cell nucleus. Cytoplasmic wild-type p53 was also found in human cervical carcinoma cell lines. (Moll, U.M. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7262-7266; and Liang, X.H. et al. (1993) *Oncogene* 8:2645-2652.)

#### Environmental Responses

Organisms respond to the environment by a number of pathways. Heat shock proteins, including hsp 70, hsp60, hsp90, and hsp 40, assist organisms in coping with heat damage to cellular proteins.

Aquaporins (AQP) are channels that transport water and, in some cases, nonionic small solutes such as urea and glycerol. Water movement is important for a number of physiological processes including renal fluid filtration, aqueous humor generation in the eye, cerebrospinal fluid production in the brain, and appropriate hydration of the lung. Aquaporins are members of the major intrinsic protein (MIP) family of membrane transporters (King, L.S. and P. Agre (1996) *Annu. Rev.*

Physiol. 58:619-648; Ishibashi, K. et al. (1997) J. Biol. Chem. 272:20782-20786). The study of aquaporins may have relevance to understanding edema formation and fluid balance in both normal physiology and disease states (King, *supra*). Mutations in AQP2 cause autosomal recessive nephrogenic diabetes insipidus (OMIM \*107777 Aquaporin 2; AQP2). Reduced AQP4 expression in skeletal muscle may be associated with Duchenne muscular dystrophy (Frigeri, A. et al. (1998) J. Clin. Invest. 102:695-703). Mutations in AQP0 cause autosomal dominant cataracts in the mouse (OMIM \*154050 Major Intrinsic Protein of Lens Fiber; MIP).

The metallothioneins (MTs) are a group of small (61 amino acids), cysteine-rich proteins that bind heavy metals such as cadmium, zinc, mercury, lead, and copper and are thought to play a role in metal detoxification or the metabolism and homeostasis of metals. Arsenite-resistance proteins have been identified in hamsters that are resistant to toxic levels of arsenite (Rossman, T.G. et al. (1997) Mutat. Res. 386:307-314).

Humans respond to light and odors by specific protein pathways. Proteins involved in light perception include rhodopsin, transducin, and cGMP phosphodiesterase. Proteins involved in odor perception include multiple olfactory receptors. Other proteins are important in human Circadian rhythms and responses to wounds.

#### Immunity and Host Defense

All vertebrates have developed sophisticated and complex immune systems that provide protection from viral, bacterial, fungal and parasitic infections. Included in these systems are the processes of humoral immunity, the complement cascade and the inflammatory response (Paul, W.E. (1993) *Fundamental Immunology*, Raven Press, Ltd., New York NY, pp.1-20).

The cellular components of the humoral immune system include six different types of leukocytes: monocytes, lymphocytes, polymorphonuclear granulocytes (consisting of neutrophils, eosinophils, and basophils) and plasma cells. Additionally, fragments of megakaryocytes, a seventh type of white blood cell in the bone marrow, occur in large numbers in the blood as platelets.

Leukocytes are formed from two stem cell lineages in bone marrow. The myeloid stem cell line produces granulocytes and monocytes and, the lymphoid stem cell produces lymphocytes. Lymphoid cells travel to the thymus, spleen and lymph nodes, where they mature and differentiate into lymphocytes. Leukocytes are responsible for defending the body against invading pathogens. Neutrophils and monocytes attack invading bacteria, viruses, and other pathogens and destroy them by phagocytosis. Monocytes enter tissues and differentiate into macrophages which are extremely phagocytic. Lymphocytes and plasma cells are a part of the immune system which recognizes specific foreign molecules and organisms and inactivates them, as well as signals other cells to attack the invaders.

Granulocytes and monocytes are formed and stored in the bone marrow until needed. Megakaryocytes are produced in bone marrow, where they fragment into platelets and are released

into the bloodstream. The main function of platelets is to activate the blood clotting mechanism. Lymphocytes and plasma cells are produced in various lymphogenous organs, including the lymph nodes, spleen, thymus, and tonsils.

Both neutrophils and macrophages exhibit chemotaxis towards sites of inflammation. Tissue  
5 inflammation in response to pathogen invasion results in production of chemo-attractants for leukocytes, such as endotoxins or other bacterial products, prostaglandins, and products of leukocytes or platelets.

Basophils participate in the release of the chemicals involved in the inflammatory process. The main function of basophils is secretion of these chemicals to such a degree that they have been  
10 referred to as "unicellular endocrine glands." A distinct aspect of basophilic secretion is that the contents of granules go directly into the extracellular environment, not into vacuoles as occurs with neutrophils, eosinophils and monocytes. Basophils have receptors for the Fc fragment of immunoglobulin E (IgE) that are not present on other leukocytes. Crosslinking of membrane IgE with anti-IgE or other ligands triggers degranulation.

Eosinophils are bi- or multi-nucleated white blood cells which contain eosinophilic granules.  
15 Their plasma membrane is characterized by Ig receptors, particularly IgG and IgE. Generally, eosinophils are stored in the bone marrow until recruited for use at a site of inflammation or invasion. They have specific functions in parasitic infections and allergic reactions, and are thought to detoxify some of the substances released by mast cells and basophils which cause inflammation. Additionally,  
20 they phagocytize antigen-antibody complexes and further help prevent spread of the inflammation.

Macrophages are monocytes that have left the blood stream to settle in tissue. Once monocytes have migrated into tissues, they do not re-enter the bloodstream. The mononuclear phagocyte system is comprised of precursor cells in the bone marrow, monocytes in circulation, and macrophages in tissues. The system is capable of very fast and extensive phagocytosis. A  
25 macrophage may phagocytize over 100 bacteria, digest them and extrude residues, and then survive for many more months. Macrophages are also capable of ingesting large particles, including red blood cells and malarial parasites. They increase several-fold in size and transform into macrophages that are characteristic of the tissue they have entered, surviving in tissues for several months.

Mononuclear phagocytes are essential in defending the body against invasion by foreign  
30 pathogens, particularly intracellular microorganisms such as M. tuberculosis, listeria, leishmania and toxoplasma. Macrophages can also control the growth of tumorous cells, via both phagocytosis and secretion of hydrolytic enzymes. Another important function of macrophages is that of processing antigen and presenting them in a biochemically modified form to lymphocytes.

The immune system responds to invading microorganisms in two major ways: antibody  
35 production and cell mediated responses. Antibodies are immunoglobulin proteins produced by B-lymphocytes which bind to specific antigens and cause inactivation or promote destruction of the

antigen by other cells. Cell-mediated immune responses involve T-lymphocytes (T cells) that react with foreign antigen on the surface of infected host cells. Depending on the type of T cell, the infected cell is either killed or signals are secreted which activate macrophages and other cells to destroy the infected cell (Paul, *supra*).

5 T-lymphocytes originate in the bone marrow or liver in fetuses. Precursor cells migrate via the blood to the thymus, where they are processed to mature into T-lymphocytes. This processing is crucial because of positive and negative selection of T cells that will react with foreign antigen and not with self molecules. After processing, T cells continuously circulate in the blood and secondary lymphoid tissues, such as lymph nodes, spleen, certain epithelium-associated tissues in the  
10 gastrointestinal tract, respiratory tract and skin. When T-lymphocytes are presented with the complementary antigen, they are stimulated to proliferate and release large numbers of activated T cells into the lymph system and the blood system. These activated T cells can survive and circulate for several days. At the same time, T memory cells are created, which remain in the lymphoid tissue for months or years. Upon subsequent exposure to that specific antigen, these memory cells will  
15 respond more rapidly and with a stronger response than induced by the original antigen. This creates an "immunological memory" that can provide immunity for years.

There are two major types of T cells: cytotoxic T cells destroy infected host cells, and helper T cells activate other white blood cells via chemical signals. One class of helper cell, T<sub>H</sub>1, activates macrophages to destroy ingested microorganisms, while another, T<sub>H</sub>2, stimulates the production of  
20 antibodies by B cells.

Cytotoxic T cells directly attack the infected target cell. In virus-infected cells, peptides derived from viral proteins are generated by the proteasome. These peptides are transported into the ER by the transporter associated with antigen processing (TAP) (Pamer, E. and P. Cresswell (1998) *Annu. Rev. Immunol.* 16:323-358). Once inside the ER, the peptides bind MHC I chains, and the  
25 peptide/MHC I complex is transported to the cell surface. Receptors on the surface of T cells bind to antigen presented on cell surface MHC molecules. Once activated by binding to antigen, T cells secrete  $\gamma$ -interferon, a signal molecule that induces the expression of genes necessary for presenting viral (or other) antigens to cytotoxic T cells. Cytotoxic T cells kill the infected cell by stimulating programmed cell death.

30 Helper T cells constitute up to 75% of the total T cell population. They regulate the immune functions by producing a variety of lymphokines that act on other cells in the immune system and on bone marrow. Among these lymphokines are: interleukins-2,3,4,5,6; granulocyte-monocyte colony stimulating factor, and  $\gamma$ -interferon.

Helper T cells are required for most B cells to respond to antigen. When an activated helper  
35 cell contacts a B cell, its centrosome and Golgi apparatus become oriented toward the B cell, aiding the directing of signal molecules, such as transmembrane-bound protein called CD40 ligand, onto the

B cell surface to interact with the CD40 transmembrane protein. Secreted signals also help B cells to proliferate and mature and, in some cases, to switch the class of antibody being produced.

B-lymphocytes (B cells) produce antibodies which react with specific antigenic proteins presented by pathogens. Once activated, B cells become filled with extensive rough endoplasmic reticulum and are known as plasma cells. As with T cells, interaction of B cells with antigen stimulates proliferation of only those B cells which produce antibody specific to that antigen. There are five classes of antibodies, known as immunoglobulins, which together comprise about 20% of total plasma protein. Each class mediates a characteristic biological response after antigen binding. Upon activation by specific antigen B cells switch from making membrane-bound antibody to secretion of that antibody.

Antibodies, or immunoglobulins (Ig), are the founding members of the Ig superfamily and the central components of the humoral immune response. Antibodies are either expressed on the surface of B cells or secreted by B cells into the circulation. Antibodies bind and neutralize blood-borne foreign antigens. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ , and  $\mu$  H-chain types. There are two types of L-chains,  $\kappa$  and  $\lambda$ , either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. Both H-chains and L-chains contain repeated Ig domains. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region. In addition, H chains such as  $\mu$  have been shown to associate with other polypeptides during differentiation of the B cell.

Antibodies can be described in terms of their two main functional domains. Antigen recognition is mediated by the Fab (antigen binding fragment) region of the antibody, while effector functions are mediated by the Fc (crystallizable fragment) region. Binding of antibody to an antigen, such as a bacterium, triggers the destruction of the antigen by phagocytic white blood cells such as macrophages and neutrophils. These cells express surface receptors that specifically bind to the antibody Fc region and allow the phagocytic cells to engulf, ingest, and degrade the antibody-bound antigen. The Fc receptors expressed by phagocytic cells are single-pass transmembrane glycoproteins of about 300 to 400 amino acids (Sears, D.W. et al. (1990) J. Immunol. 144:371-378).

The extracellular portion of the Fc receptor typically contains two or three Ig domains.

Diseases which cause over- or under-abundance of any one type of leukocyte usually result in the entire immune defense system becoming involved. A well-known autoimmune disease is AIDS (Acquired Immunodeficiency Syndrome) where the number of helper T cells is depleted, leaving the patient susceptible to infection by microorganisms and parasites. Another widespread medical condition attributable to the immune system is that of allergic reactions to certain antigens. Allergic reactions include: hay fever, asthma, anaphylaxis, and urticaria (hives). Leukemias are an excess production of white blood cells, to the point where a major portion of the body's metabolic resources are directed solely at proliferation of white blood cells, leaving other tissues to starve. Leukopenia or agranulocytosis occurs when the bone marrow stops producing white blood cells. This leaves the body unprotected against foreign microorganisms, including those which normally inhabit skin, mucous membranes, and gastrointestinal tract. If all white blood cell production stops completely, infection will occur within two days and death may follow only 1 to 4 days later.

Impaired phagocytosis occurs in several diseases, including monocytic leukemia, systemic lupus, and granulomatous disease. In such a situation, macrophages can phagocytize normally, but the enveloped organism is not killed. A defect in the plasma membrane enzyme which converts oxygen to lethally reactive forms results in abscess formation in liver, lungs, spleen, lymph nodes, and beneath the skin. Eosinophilia is an excess of eosinophils commonly observed in patients with allergies (hay fever, asthma), allergic reactions to drugs, rheumatoid arthritis, and cancers (Hodgkin's disease, lung, and liver cancer) (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, Inc., New York NY).

Host defense is further augmented by the complement system. The complement system serves as an effector system and is involved in infectious agent recognition. It can function as an independent immune network or in conjunction with other humoral immune responses. The complement system is comprised of numerous plasma and membrane proteins that act in a cascade of reaction sequences whereby one component activates the next. The result is a rapid and amplified response to infection through either an inflammatory response or increased phagocytosis.

The complement system has more than 30 protein components which can be divided into functional groupings including modified serine proteases, membrane-binding proteins and regulators of complement activation. Activation occurs through two different pathways the classical and the alternative. Both pathways serve to destroy infectious agents through distinct triggering mechanisms that eventually merge with the involvement of the component C3.

The classical pathway requires antibody binding to infectious agent antigens. The antibodies serve to define the target and initiate the complement system cascade, culminating in the destruction of the infectious agent. In this pathway, since the antibody guides initiation of the process, the complement can be seen as an effector arm of the humoral immune system.

The alternative pathway of the complement system does not require the presence of pre-existing antibodies for targeting infectious agent destruction. Rather, this pathway, through low levels of an activated component, remains constantly primed and provides surveillance in the non-immune host to enable targeting and destruction of infectious agents. In this case foreign material triggers the cascade, thereby facilitating phagocytosis or lysis (Paul, *supra*, pp.918-919).

Another important component of host defense is the process of inflammation. Inflammatory responses are divided into four categories on the basis of pathology and include allergic inflammation, cytotoxic antibody mediated inflammation, immune complex mediated inflammation and monocyte mediated inflammation. Inflammation manifests as a combination of each of these forms with one predominating.

Allergic acute inflammation is observed in individuals wherein specific antigens stimulate IgE antibody production. Mast cells and basophils are subsequently activated by the attachment of antigen-IgE complexes, resulting in the release of cytoplasmic granule contents such as histamine. The products of activated mast cells can increase vascular permeability and constrict the smooth muscle of breathing passages, resulting in anaphylaxis or asthma. Acute inflammation is also mediated by cytotoxic antibodies and can result in the destruction of tissue through the binding of complement-fixing antibodies to cells. The responsible antibodies are of the IgG or IgM types. Resultant clinical disorders include autoimmune hemolytic anemia and thrombocytopenia as associated with systemic lupus erythematosus.

Immune complex mediated acute inflammation involves the IgG or IgM antibody types which combine with antigen to activate the complement cascade. When such immune complexes bind to neutrophils and macrophages they activate the respiratory burst to form protein- and vessel-damaging agents such as hydrogen peroxide, hydroxyl radical, hypochlorous acid, and chloramines. Clinical manifestations include rheumatoid arthritis and systemic lupus erythematosus.

In chronic inflammation or delayed-type hypersensitivity, macrophages are activated and process antigen for presentation to T cells that subsequently produce lymphokines and monokines. This type of inflammatory response is likely important for defense against intracellular parasites and certain viruses. Clinical associations include, granulomatous disease, tuberculosis, leprosy, and sarcoidosis (Paul, W.E., *supra*, pp.1017-1018).

### **Extracellular Information Transmission Molecules**

Intercellular communication is essential for the growth and survival of multicellular organisms, and in particular, for the function of the endocrine, nervous, and immune systems. In addition, intercellular communication is critical for developmental processes such as tissue construction and organogenesis, in which cell proliferation, cell differentiation, and morphogenesis must be spatially and temporally regulated in a precise and coordinated manner. Cells communicate



with one another through the secretion and uptake of diverse types of signaling molecules such as hormones, growth factors, neuropeptides, and cytokines.

### Hormones

Hormones are signaling molecules that coordinately regulate basic physiological processes from embryogenesis throughout adulthood. These processes include metabolism, respiration, reproduction, excretion, fetal tissue differentiation and organogenesis, growth and development, homeostasis, and the stress response. Hormonal secretions and the nervous system are tightly integrated and interdependent. Hormones are secreted by endocrine glands, primarily the hypothalamus and pituitary, the thyroid and parathyroid, the pancreas, the adrenal glands, and the ovaries and testes.

The secretion of hormones into the circulation is tightly controlled. Hormones are often secreted in diurnal, pulsatile, and cyclic patterns. Hormone secretion is regulated by perturbations in blood biochemistry, by other upstream-acting hormones, by neural impulses, and by negative feedback loops. Blood hormone concentrations are constantly monitored and adjusted to maintain optimal, steady-state levels. Once secreted, hormones act only on those target cells that express specific receptors.

Most disorders of the endocrine system are caused by either hyposecretion or hypersecretion of hormones. Hyposecretion often occurs when a hormone's gland of origin is damaged or otherwise impaired. Hypersecretion often results from the proliferation of tumors derived from hormone-secreting cells. Inappropriate hormone levels may also be caused by defects in regulatory feedback loops or in the processing of hormone precursors. Endocrine malfunction may also occur when the target cell fails to respond to the hormone.

Hormones can be classified biochemically as polypeptides, steroids, eicosanoids, or amines. Polypeptides, which include diverse hormones such as insulin and growth hormone, vary in size and function and are often synthesized as inactive precursors that are processed intracellularly into mature, active forms. Amines, which include epinephrine and dopamine, are amino acid derivatives that function in neuroendocrine signaling. Steroids, which include the cholesterol-derived hormones estrogen and testosterone, function in sexual development and reproduction. Eicosanoids, which include prostaglandins and prostacyclins, are fatty acid derivatives that function in a variety of processes. Most polypeptides and some amines are soluble in the circulation where they are highly susceptible to proteolytic degradation within seconds after their secretion. Steroids and lipids are insoluble and must be transported in the circulation by carrier proteins. The following discussion will focus primarily on polypeptide hormones.

Hormones secreted by the hypothalamus and pituitary gland play a critical role in endocrine function by coordinately regulating hormonal secretions from other endocrine glands in response to neural signals. Hypothalamic hormones include thyrotropin-releasing hormone, gonadotropin-

releasing hormone, somatostatin, growth-hormone releasing factor, corticotropin-releasing hormone, substance P, dopamine, and prolactin-releasing hormone. These hormones directly regulate the secretion of hormones from the anterior lobe of the pituitary. Hormones secreted by the anterior pituitary include adrenocorticotrophic hormone (ACTH), melanocyte-stimulating hormone, 5 somatotrophic hormones such as growth hormone and prolactin, glycoprotein hormones such as thyroid-stimulating hormone, luteinizing hormone (LH), and follicle-stimulating hormone (FSH),  $\beta$ -lipotropin, and  $\beta$ -endorphins. These hormones regulate hormonal secretions from the thyroid, pancreas, and adrenal glands, and act directly on the reproductive organs to stimulate ovulation and spermatogenesis. The posterior pituitary synthesizes and secretes antidiuretic hormone (ADH, 10 vasopressin) and oxytocin.

Disorders of the hypothalamus and pituitary often result from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma. Such disorders have profound effects on the function of other endocrine glands. Disorders 15 associated with hypopituitarism include hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism. Disorders associated with hyperpituitarism include acromegaly, gigantism, and syndrome of inappropriate ADH secretion (SIADH), often caused by benign adenomas.

Hormones secreted by the thyroid and parathyroid primarily control metabolic rates and the 20 regulation of serum calcium levels, respectively. Thyroid hormones include calcitonin, somatostatin, and thyroid hormone. The parathyroid secretes parathyroid hormone. Disorders associated with hypothyroidism include goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism. Disorders associated with hyperthyroidism include thyrotoxicosis and its various 25 forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease. Disorders associated with hyperparathyroidism include Conn disease (chronic hypercalcemia) leading to bone resorption and parathyroid hyperplasia.

Hormones secreted by the pancreas regulate blood glucose levels by modulating the rates of carbohydrate, fat, and protein metabolism. Pancreatic hormones include insulin, glucagon, amylin,  $\gamma$ - 30 aminobutyric acid, gastrin, somatostatin, and pancreatic polypeptide. The principal disorder associated with pancreatic dysfunction is diabetes mellitus caused by insufficient insulin activity. Diabetes mellitus is generally classified as either Type I (insulin-dependent, juvenile diabetes) or Type II (non-insulin-dependent, adult diabetes). The treatment of both forms by insulin replacement therapy is well known. Diabetes mellitus often leads to acute complications such as hypoglycemia 35 (insulin shock), coma, diabetic ketoacidosis, lactic acidosis, and chronic complications leading to disorders of the eye, kidney, skin, bone, joint, cardiovascular system, nervous system, and to

decreased resistance to infection.

The anatomy, physiology, and diseases related to hormonal function are reviewed in McCance, K.L. and S.E. Huether (1994) Pathophysiology: The Biological Basis for Disease in Adults and Children, Mosby-Year Book, Inc., St. Louis MO; Greenspan, F.S. and J.D. Baxter (1994) Basic and Clinical Endocrinology, Appleton and Lange, East Norwalk CT.

#### Growth Factors

Growth factors are secreted proteins that mediate intercellular communication. Unlike hormones, which travel great distances via the circulatory system, most growth factors are primarily local mediators that act on neighboring cells. Most growth factors contain a hydrophobic N-terminal signal peptide sequence which directs the growth factor into the secretory pathway. Most growth factors also undergo post-translational modifications within the secretory pathway. These modifications can include proteolysis, glycosylation, phosphorylation, and intramolecular disulfide bond formation. Once secreted, growth factors bind to specific receptors on the surfaces of neighboring target cells, and the bound receptors trigger intracellular signal transduction pathways. These signal transduction pathways elicit specific cellular responses in the target cells. These responses can include the modulation of gene expression and the stimulation or inhibition of cell division, cell differentiation, and cell motility.

Growth factors fall into at least two broad and overlapping classes. The broadest class includes the large polypeptide growth factors, which are wide-ranging in their effects. These factors include epidermal growth factor (EGF), fibroblast growth factor (FGF), transforming growth factor- $\beta$  (TGF- $\beta$ ), insulin-like growth factor (IGF), nerve growth factor (NGF), and platelet-derived growth factor (PDGF), each defining a family of numerous related factors. The large polypeptide growth factors, with the exception of NGF, act as mitogens on diverse cell types to stimulate wound healing, bone synthesis and remodeling, extracellular matrix synthesis, and proliferation of epithelial, epidermal, and connective tissues. Members of the TGF- $\beta$ , EGF, and FGF families also function as inductive signals in the differentiation of embryonic tissue. NGF functions specifically as a neurotrophic factor, promoting neuronal growth and differentiation.

Another class of growth factors includes the hematopoietic growth factors, which are narrow in their target specificity. These factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. These factors include the colony-stimulating factors (G-CSF, M-CSF, GM-CSF, and CSF1-3), erythropoietin, and the cytokines. The cytokines are specialized hematopoietic factors secreted by cells of the immune system and are discussed in detail below.

Growth factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Overexpression of the large polypeptide growth factors promotes the

proliferation and transformation of cells in culture. Inappropriate expression of these growth factors by tumor cells in vivo may contribute to tumor vascularization and metastasis. Inappropriate activity of hematopoietic growth factors can result in anemias, leukemias, and lymphomas. Moreover, growth factors are both structurally and functionally related to oncoproteins, the potentially cancer-

5 causing products of proto-oncogenes. Certain FGF and PDGF family members are themselves homologous to oncoproteins, whereas receptors for some members of the EGF, NGF, and FGF families are encoded by proto-oncogenes. Growth factors also affect the transcriptional regulation of both proto-oncogenes and oncosuppressor genes (Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor MI; McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical

10 Approach, Oxford University Press, New York NY; Habenicht, A., ed. (1990) Growth Factors, Differentiation Factors, and Cytokines, Springer-Verlag, New York NY).

In addition, some of the large polypeptide growth factors play crucial roles in the induction of the primordial germ layers in the developing embryo. This induction ultimately results in the formation of the embryonic mesoderm, ectoderm, and endoderm which in turn provide the framework

15 for the entire adult body plan. Disruption of this inductive process would be catastrophic to embryonic development.

#### Small Peptide Factors - Neuropeptides and Vasomediators

Neuropeptides and vasomediators (NP/VM) comprise a family of small peptide factors, typically of 20 amino acids or less. These factors generally function in neuronal excitation and

20 inhibition of vasoconstriction/vasodilation, muscle contraction, and hormonal secretions from the brain and other endocrine tissues. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling

25 molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin, gastrin, and many of the peptide hormones discussed above. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in signaling cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford

30 University Press, New York NY, pp. 57-62.)

#### Cytokines

Cytokines comprise a family of signaling molecules that modulate the immune system and the inflammatory response. Cytokines are usually secreted by leukocytes, or white blood cells, in response to injury or infection. Cytokines function as growth and differentiation factors that act

35 primarily on cells of the immune system such as B- and T-lymphocytes, monocytes, macrophages, and granulocytes. Like other signaling molecules, cytokines bind to specific plasma membrane

receptors and trigger intracellular signal transduction pathways which alter gene expression patterns. There is considerable potential for the use of cytokines in the treatment of inflammation and immune system disorders.

Cytokine structure and function have been extensively characterized in vitro. Most cytokines  
5 are small polypeptides of about 30 kilodaltons or less. Over 50 cytokines have been identified from human and rodent sources. Examples of cytokine subfamilies include the interferons (IFN- $\alpha$ , - $\beta$ , and - $\gamma$ ), the interleukins (IL1-IL13), the tumor necrosis factors (TNF- $\alpha$  and - $\beta$ ), and the chemokines. Many cytokines have been produced using recombinant DNA techniques, and the activities of individual cytokines have been determined in vitro. These activities include regulation of leukocyte  
10 proliferation, differentiation, and motility.

The activity of an individual cytokine in vitro may not reflect the full scope of that cytokine's activity in vivo. Cytokines are not expressed individually in vivo but are instead expressed in combination with a multitude of other cytokines when the organism is challenged with a stimulus. Together, these cytokines collectively modulate the immune response in a manner appropriate for that  
15 particular stimulus. Therefore, the physiological activity of a cytokine is determined by the stimulus itself and by complex interactive networks among co-expressed cytokines which may demonstrate both synergistic and antagonistic relationships.

Chemokines comprise a cytokine subfamily with over 30 members. (Reviewed in Wells, T. N.C. and M.C. Peitsch (1997) *J. Leukoc. Biol.* 61:545-550.) Chemokines were initially identified as  
20 chemotactic proteins that recruit monocytes and macrophages to sites of inflammation. Recent evidence indicates that chemokines may also play key roles in hematopoiesis and HIV-1 infection. Chemokines are small proteins which range from about 6-15 kilodaltons in molecular weight. Chemokines are further classified as C, CC, CXC, or CX<sub>3</sub>C based on the number and position of critical cysteine residues. The CC chemokines, for example, each contain a conserved motif  
25 consisting of two consecutive cysteines followed by two additional cysteines which occur downstream at 24- and 16-residue intervals, respectively (ExpASY PROSITE database, documents PS00472 and PDOC00434). The presence and spacing of these four cysteine residues are highly conserved, whereas the intervening residues diverge significantly. However, a conserved tyrosine located about 15 residues downstream of the cysteine doublet seems to be important for chemotactic  
30 activity. Most of the human genes encoding CC chemokines are clustered on chromosome 17, although there are a few examples of CC chemokine genes that map elsewhere. Other chemokines include lymphotactin (C chemokine); macrophage chemotactic and activating factor (MCAF/MCP-1; CC chemokine); platelet factor 4 and IL-8 (CXC chemokines); and fractalkine and neurotractin (CX<sub>3</sub>C chemokines). (Reviewed in Luster, A.D. (1998) *N. Engl. J. Med.* 338:436-445.)

35

## Receptor Molecules

The term receptor describes proteins that specifically recognize other molecules. The category is broad and includes proteins with a variety of functions. The bulk of receptors are cell surface proteins which bind extracellular ligands and produce cellular responses in the areas of growth, differentiation, endocytosis, and immune response. Other receptors facilitate the selective transport of proteins out of the endoplasmic reticulum and localize enzymes to particular locations in the cell. The term may also be applied to proteins which act as receptors for ligands with known or unknown chemical composition and which interact with other cellular components. For example, the steroid hormone receptors bind to and regulate transcription of DNA.

Regulation of cell proliferation, differentiation, and migration is important for the formation and function of tissues. Regulatory proteins such as growth factors coordinately control these cellular processes and act as mediators in cell-cell signaling pathways. Growth factors are secreted proteins that bind to specific cell-surface receptors on target cells. The bound receptors trigger intracellular signal transduction pathways which activate various downstream effectors that regulate gene expression, cell division, cell differentiation, cell motility, and other cellular processes.

Cell surface receptors are typically integral plasma membrane proteins. These receptors recognize hormones such as catecholamines; peptide hormones; growth and differentiation factors; small peptide factors such as thyrotropin-releasing hormone; galanin, somatostatin, and tachykinins; and circulatory system-borne signaling molecules. Cell surface receptors on immune system cells recognize antigens, antibodies, and major histocompatibility complex (MHC)-bound peptides. Other cell surface receptors bind ligands to be internalized by the cell. This receptor-mediated endocytosis functions in the uptake of low density lipoproteins (LDL), transferrin, glucose- or mannose-terminal glycoproteins, galactose-terminal glycoproteins, immunoglobulins, phosphovitellogenins, fibrin, proteinase-inhibitor complexes, plasminogen activators, and thrombospondin (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, p. 723; Mikhailenko, I. et al. (1997) J. Biol. Chem. 272:6784-6791).

#### Receptor Protein Kinases

Many growth factor receptors, including receptors for epidermal growth factor, platelet-derived growth factor, fibroblast growth factor, as well as the growth modulator  $\alpha$ -thrombin, contain intrinsic protein kinase activities. When growth factor binds to the receptor, it triggers the autophosphorylation of a serine, threonine, or tyrosine residue on the receptor. These phosphorylated sites are recognition sites for the binding of other cytoplasmic signaling proteins. These proteins participate in signaling pathways that eventually link the initial receptor activation at the cell surface to the activation of a specific intracellular target molecule. In the case of tyrosine residue autophosphorylation, these signaling proteins contain a common domain referred to as a Src homology (SH) domain. SH2 domains and SH3 domains are found in phospholipase C- $\gamma$ , PI-3-K p85 regulatory subunit, Ras-GTPase activating protein, and pp60<sup>c-src</sup> (Lowenstein, E.J. et al. (1992) Cell

70:431-442). The cytokine family of receptors share a different common binding domain and include transmembrane receptors for growth hormone (GH), interleukins, erythropoietin, and prolactin.

Other receptors and second messenger-binding proteins have intrinsic serine/threonine protein kinase activity. These include activin/TGF- $\beta$ /BMP-superfamily receptors, calcium- and  
 5 diacylglycerol-activated/phospholipid-dependant protein kinase (PK-C), and RNA-dependant protein kinase (PK-R). In addition, other serine/threonine protein kinases, including nematode Twitchin, have fibronectin-like, immunoglobulin C2-like domains.

#### G-Protein Coupled Receptors

G-protein coupled receptors (GPCRs) are integral membrane proteins characterized by the  
 10 presence of seven hydrophobic transmembrane domains which span the plasma membrane and form a bundle of antiparallel alpha ( $\alpha$ ) helices. These proteins range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) Eur. J. Biochem. 196:1-10; Coughlin, S.R. (1994) Curr. Opin. Cell Biol. 6:191-197). The amino-terminus of the GPCR is extracellular, of variable length and often glycosylated; the carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops  
 15 of the GPCR alternate with intracellular loops and link the transmembrane domains. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. The transmembrane domains account for structural and functional features of the receptor. In most cases, the bundle of  $\alpha$  helices forms a binding pocket. In addition, the extracellular N-terminal segment or one or more of the three extracellular loops may also participate in ligand binding.  
 20 Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. The activated receptor, in turn, interacts with an intracellular heterotrimeric guanine nucleotide binding (G) protein complex which mediates further intracellular signaling activities, generally the production of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate, or interactions with ion channel proteins (Baldwin, J.M. (1994) Curr. Opin. Cell  
 25 Biol. 6:180-190).

GPCRs include those for acetylcholine, adenosine, epinephrine and norepinephrine, bombesin, bradykinin, chemokines, dopamine, endothelin,  $\gamma$ -aminobutyric acid (GABA), follicle-stimulating hormone (FSH), glutamate, gonadotropin-releasing hormone (GnRH), hepatocyte growth factor, histamine, leukotrienes, melanocortins, neuropeptide Y, opioid peptides, opsins, prostanoids,  
 30 serotonin, somatostatin, tachykinins, thrombin, thyrotropin-releasing hormone (TRH), vasoactive intestinal polypeptide family, vasopressin and oxytocin, and orphan receptors.

GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, *supra*). For instance, retinitis pigmentosa may arise from mutations in the rhodopsin gene. Rhodopsin is the retinal photoreceptor which is located  
 35 within the discs of the eye rod cell. Parma, J. et al. (1993, Nature 365:649-651) report that somatic activating mutations in the thyrotropin receptor cause hyperfunctioning thyroid adenomas and suggest

that certain GPCRs susceptible to constitutive activation may behave as protooncogenes.

#### Nuclear Receptors

Nuclear receptors bind small molecules such as hormones or second messengers, leading to increased receptor-binding affinity to specific chromosomal DNA elements. In addition the affinity  
5 for other nuclear proteins may also be altered. Such binding and protein-protein interactions may regulate and modulate gene expression. Examples of such receptors include the steroid hormone receptors family, the retinoic acid receptors family, and the thyroid hormone receptors family.

#### Ligand-Gated Receptor Ion Channels

Ligand-gated receptor ion channels fall into two categories. The first category, extracellular  
10 ligand-gated receptor ion channels (ELGs), rapidly transduce neurotransmitter-binding events into electrical signals, such as fast synaptic neurotransmission. ELG function is regulated by post-translational modification. The second category, intracellular ligand-gated receptor ion channels (ILGs), are activated by many intracellular second messengers and do not require post-translational modification(s) to effect a channel-opening response.

15 ELGs depolarize excitable cells to the threshold of action potential generation. In non-excitable cells, ELGs permit a limited calcium ion-influx during the presence of agonist. ELGs include channels directly gated by neurotransmitters such as acetylcholine, L-glutamate, glycine, ATP, serotonin, GABA, and histamine. ELG genes encode proteins having strong structural and functional similarities. ILGs are encoded by distinct and unrelated gene families and include  
20 receptors for cAMP, cGMP, calcium ions, ATP, and metabolites of arachidonic acid.

#### Macrophage Scavenger Receptors

Macrophage scavenger receptors with broad ligand specificity may participate in the binding of low density lipoproteins (LDL) and foreign antigens. Scavenger receptors types I and II are  
25 trimeric membrane proteins with each subunit containing a small N-terminal intracellular domain, a transmembrane domain, a large extracellular domain, and a C-terminal cysteine-rich domain. The extracellular domain contains a short spacer domain, an  $\alpha$ -helical coiled-coil domain, and a triple helical collagenous domain. These receptors have been shown to bind a spectrum of ligands, including chemically modified lipoproteins and albumin, polyribonucleotides, polysaccharides, phospholipids, and asbestos (Matsumoto, A. et al. (1990) Proc. Natl. Acad. Sci. USA 87:9133-9137;  
30 Elomaa, O. et al. (1995) Cell 80:603-609). The scavenger receptors are thought to play a key role in atherogenesis by mediating uptake of modified LDL in arterial walls, and in host defense by binding bacterial endotoxins, bacteria, and protozoa.

#### T-Cell Receptors

T cells play a dual role in the immune system as effectors and regulators, coupling antigen  
35 recognition with the transmission of signals that induce cell death in infected cells and stimulate proliferation of other immune cells. Although a population of T cells can recognize a wide range of



different antigens, an individual T cell can only recognize a single antigen and only when it is presented to the T cell receptor (TCR) as a peptide complexed with a major histocompatibility molecule (MHC) on the surface of an antigen presenting cell. The TCR on most T cells consists of immunoglobulin-like integral membrane glycoproteins containing two polypeptide subunits,  $\alpha$  and  $\beta$ , of similar molecular weight. Both TCR subunits have an extracellular domain containing both variable and constant regions, a transmembrane domain that traverses the membrane once, and a short intracellular domain (Saito, H. et al. (1984) *Nature* 309:757-762). The genes for the TCR subunits are constructed through somatic rearrangement of different gene segments. Interaction of antigen in the proper MHC context with the TCR initiates signaling cascades that induce the proliferation, maturation, and function of cellular components of the immune system (Weiss, A. (1991) *Annu. Rev. Genet.* 25:487-510). Rearrangements in TCR genes and alterations in TCR expression have been noted in lymphomas, leukemias, autoimmune disorders, and immunodeficiency disorders (Aisenberg, A.C. et al. (1985) *N. Engl. J. Med.* 313:529-533; Weiss, *supra*).

## **Intracellular Signaling Molecules**

Intracellular signaling is the general process by which cells respond to extracellular signals (hormones, neurotransmitters, growth and differentiation factors, etc.) through a cascade of biochemical reactions that begins with the binding of a signaling molecule to a cell membrane receptor and ends with the activation of an intracellular target molecule. Intermediate steps in the process involve the activation of various cytoplasmic proteins by phosphorylation via protein kinases, and their deactivation by protein phosphatases, and the eventual translocation of some of these activated proteins to the cell nucleus where the transcription of specific genes is triggered. The intracellular signaling process regulates all types of cell functions including cell proliferation, cell differentiation, and gene transcription, and involves a diversity of molecules including protein kinases and phosphatases, and second messenger molecules, such as cyclic nucleotides, calcium-calmodulin, inositol, and various mitogens, that regulate protein phosphorylation.

### Protein Phosphorylation

Protein kinases and phosphatases play a key role in the intracellular signaling process by controlling the phosphorylation and activation of various signaling proteins. The high energy phosphate for this reaction is generally transferred from the adenosine triphosphate molecule (ATP) to a particular protein by a protein kinase and removed from that protein by a protein phosphatase. Protein kinases are roughly divided into two groups: those that phosphorylate tyrosine residues (protein tyrosine kinases, PTK) and those that phosphorylate serine or threonine residues (serine/threonine kinases, STK). A few protein kinases have dual specificity for serine/threonine and tyrosine residues. Almost all kinases contain a conserved 250-300 amino acid catalytic domain containing specific residues and sequence motifs characteristic of the kinase family (Hardie, G. and

S. Hanks (1995) The Protein Kinase Facts Books, Vol I:7-20, Academic Press, San Diego CA).

STKs include the second messenger dependent protein kinases such as the cyclic-AMP dependent protein kinases (PKA), involved in mediating hormone-induced cellular responses; calcium-calmodulin (CaM) dependent protein kinases, involved in regulation of smooth muscle contraction, glycogen breakdown, and neurotransmission; and the mitogen-activated protein kinases (MAP) which mediate signal transduction from the cell surface to the nucleus via phosphorylation cascades. Altered PKA expression is implicated in a variety of disorders and diseases including cancer, thyroid disorders, diabetes, atherosclerosis, and cardiovascular disease (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York NY, pp. 416-431, 1887).

PTKs are divided into transmembrane, receptor PTKs and nontransmembrane, non-receptor PTKs. Transmembrane PTKs are receptors for most growth factors. Non-receptor PTKs lack transmembrane regions and, instead, form complexes with the intracellular regions of cell surface receptors. Receptors that function through non-receptor PTKs include those for cytokines and hormones (growth hormone and prolactin) and antigen-specific receptors on T and B lymphocytes. Many of these PTKs were first identified as the products of mutant oncogenes in cancer cells in which their activation was no longer subject to normal cellular controls. In fact, about one third of the known oncogenes encode PTKs, and it is well known that cellular transformation (oncogenesis) is often accompanied by increased tyrosine phosphorylation activity (Charbonneau, H. and N.K. Tonks (1992) *Annu. Rev. Cell Biol.* 8:463-493).

An additional family of protein kinases previously thought to exist only in procaryotes is the histidine protein kinase family (HPK). HPKs bear little homology with mammalian STKs or PTKs but have distinctive sequence motifs of their own (Davie, J.R. et al. (1995) *J. Biol. Chem.* 270:19861-19867). A histidine residue in the N-terminal half of the molecule (region I) is an autophosphorylation site. Three additional motifs located in the C-terminal half of the molecule include an invariant asparagine residue in region II and two glycine-rich loops characteristic of nucleotide binding domains in regions III and IV. Recently a branched chain alpha-ketoacid dehydrogenase kinase has been found with characteristics of HPK in rat (Davie, *supra*).

Protein phosphatases regulate the effects of protein kinases by removing phosphate groups from molecules previously activated by kinases. The two principal categories of protein phosphatases are the protein (serine/threonine) phosphatases (PPs) and the protein tyrosine phosphatases (PTPs). PPs dephosphorylate phosphoserine/threonine residues and are important regulators of many cAMP-mediated hormone responses (Cohen, P. (1989) *Annu. Rev. Biochem.* 58:453-508). PTPs reverse the effects of protein tyrosine kinases and play a significant role in cell cycle and cell signaling processes (Charbonneau, *supra*). As previously noted, many PTKs are encoded by oncogenes, and oncogenesis is often accompanied by increased tyrosine phosphorylation activity. It is therefore possible that PTPs may prevent or reverse cell transformation and the growth

of various cancers by controlling the levels of tyrosine phosphorylation in cells. This hypothesis is supported by studies showing that overexpression of PTPs can suppress transformation in cells, and that specific inhibition of PTPs can enhance cell transformation (Charbonneau, *supra*).

#### Phospholipid and Inositol-Phosphate Signaling

5 Inositol phospholipids (phosphoinositides) are involved in an intracellular signaling pathway that begins with binding of a signaling molecule to a G-protein linked receptor in the plasma membrane. This leads to the phosphorylation of phosphatidylinositol (PI) residues on the inner side of the plasma membrane to the biphosphate state ( $\text{PIP}_2$ ) by inositol kinases. Simultaneously, the G-protein linked receptor binding stimulates a trimeric G-protein which in turn activates a  
10 phosphoinositide-specific phospholipase C- $\beta$ . Phospholipase C- $\beta$  then cleaves  $\text{PIP}_2$  into two products, inositol triphosphate ( $\text{IP}_3$ ) and diacylglycerol. These two products act as mediators for separate signaling events.  $\text{IP}_3$  diffuses through the plasma membrane to induce calcium release from the endoplasmic reticulum (ER), while diacylglycerol remains in the membrane and helps activate protein kinase C, an STK that phosphorylates selected proteins in the target cell. The calcium  
15 response initiated by  $\text{IP}_3$  is terminated by the dephosphorylation of  $\text{IP}_3$  by specific inositol phosphatases. Cellular responses that are mediated by this pathway are glycogen breakdown in the liver in response to vasopressin, smooth muscle contraction in response to acetylcholine, and thrombin-induced platelet aggregation.

#### Cyclic Nucleotide Signaling

20 Cyclic nucleotides (cAMP and cGMP) function as intracellular second messengers to transduce a variety of extracellular signals including hormones, light, and neurotransmitters. In particular, cyclic-AMP dependent protein kinases (PKA) are thought to account for all of the effects of cAMP in most mammalian cells, including various hormone-induced cellular responses. Visual excitation and the phototransmission of light signals in the eye is controlled by cyclic-GMP  
25 regulated,  $\text{Ca}^{2+}$ -specific channels. Because of the importance of cellular levels of cyclic nucleotides in mediating these various responses, regulating the synthesis and breakdown of cyclic nucleotides is an important matter. Thus adenylyl cyclase, which synthesizes cAMP from AMP, is activated to increase cAMP levels in muscle by binding of adrenaline to  $\beta$ -adrenergic receptors, while activation of guanylate cyclase and increased cGMP levels in photoreceptors leads to reopening of the  
30  $\text{Ca}^{2+}$ -specific channels and recovery of the dark state in the eye. In contrast, hydrolysis of cyclic nucleotides by cAMP and cGMP-specific phosphodiesterases (PDEs) produces the opposite of these and other effects mediated by increased cyclic nucleotide levels. PDEs appear to be particularly important in the regulation of cyclic nucleotides, considering the diversity found in this family of proteins. At least seven families of mammalian PDEs (PDE1-7) have been identified based on  
35 substrate specificity and affinity, sensitivity to cofactors, and sensitivity to inhibitory drugs (Beavo, J.A. (1995) *Physiological Reviews* 75:725-748). PDE inhibitors have been found to be particularly

useful in treating various clinical disorders. Rolipram, a specific inhibitor of PDE4, has been used in the treatment of depression, and similar inhibitors are undergoing evaluation as anti-inflammatory agents. Theophylline is a nonspecific PDE inhibitor used in the treatment of bronchial asthma and other respiratory diseases (Banner, K.H. and C.P. Page (1995) Eur. Respir. J. 8:996-1000).

5 G-Protein Signaling

Guanine nucleotide binding proteins (G-proteins) are critical mediators of signal transduction between a particular class of extracellular receptors, the G-protein coupled receptors (GPCR), and intracellular second messengers such as cAMP and  $\text{Ca}^{2+}$ . G-proteins are linked to the cytosolic side of a GPCR such that activation of the GPCR by ligand binding stimulates binding of the G-protein to  
10 GTP, inducing an "active" state in the G-protein. In the active state, the G-protein acts as a signal to trigger other events in the cell such as the increase of cAMP levels or the release of  $\text{Ca}^{2+}$  into the cytosol from the ER, which, in turn, regulate phosphorylation and activation of other intracellular proteins. Recycling of the G-protein to the inactive state involves hydrolysis of the bound GTP to GDP by a GTPase activity in the G-protein. (See Alberts, B. et al. (1994) Molecular Biology of the  
15 Cell, Garland Publishing, Inc., New York NY, pp.734-759.) Two structurally distinct classes of G-proteins are recognized: heterotrimeric G-proteins, consisting of three different subunits, and monomeric, low molecular weight (LMW), G-proteins consisting of a single polypeptide chain.

The three polypeptide subunits of heterotrimeric G-proteins are the  $\alpha$ ,  $\beta$ , and  $\gamma$  subunits. The  $\alpha$  subunit binds and hydrolyzes GTP. The  $\beta$  and  $\gamma$  subunits form a tight complex that anchors the  
20 protein to the inner side of the plasma membrane. The  $\beta$  subunits, also known as G- $\beta$  proteins or  $\beta$  transducins, contain seven tandem repeats of the WD-repeat sequence motif, a motif found in many proteins with regulatory functions. Mutations and variant expression of  $\beta$  transducin proteins are linked with various disorders (Neer, E.J. et al. (1994) Nature 371:297-300; Margottin, F. et al. (1998) Mol. Cell 1:565-574).

LMW GTP-proteins are GTPases which regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. They consist of single polypeptides which, like the  $\alpha$  subunit of the heterotrimeric G-proteins, are able to bind and hydrolyze GTP, thus cycling between an inactive and an active state. At least sixty members of the LMW G-protein superfamily have been identified and are currently grouped into the six subfamilies of ras, rho, arf, sar1, ran, and rab.  
25 Activated ras genes were initially found in human cancers, and subsequent studies confirmed that ras function is critical in determining whether cells continue to grow or become differentiated. Other members of the LMW G-protein superfamily have roles in signal transduction that vary with the function of the activated genes and the locations of the G-proteins.

Guanine nucleotide exchange factors regulate the activities of LMW G-proteins by  
35 determining whether GTP or GDP is bound. GTPase-activating protein (GAP) binds to GTP-ras and

induces it to hydrolyze GTP to GDP. In contrast, guanine nucleotide releasing protein (GNRP) binds to GDP-ras and induces the release of GDP and the binding of GTP.

Other regulators of G-protein signaling (RGS) also exist that act primarily by negatively regulating the G-protein pathway by an unknown mechanism (Druey, K.M. et al. (1996) Nature 379:742-746). Some 15 members of the RGS family have been identified. RGS family members are related structurally through similarities in an approximately 120 amino acid region termed the RGS domain and functionally by their ability to inhibit the interleukin (cytokine) induction of MAP kinase in cultured mammalian 293T cells (Druey, supra).

#### Calcium Signaling Molecules

10  $\text{Ca}^{+2}$  is another second messenger molecule that is even more widely used as an intracellular mediator than cAMP. Two pathways exist by which  $\text{Ca}^{+2}$  can enter the cytosol in response to extracellular signals: One pathway acts primarily in nerve signal transduction where  $\text{Ca}^{+2}$  enters a nerve terminal through a voltage-gated  $\text{Ca}^{+2}$  channel. The second is a more ubiquitous pathway in which  $\text{Ca}^{+2}$  is released from the ER into the cytosol in response to binding of an extracellular  
15 signaling molecule to a receptor.  $\text{Ca}^{2+}$  directly activates regulatory enzymes, such as protein kinase C, which trigger signal transduction pathways.  $\text{Ca}^{2+}$  also binds to specific  $\text{Ca}^{2+}$ -binding proteins (CBPs) such as calmodulin (CaM) which then activate multiple target proteins in the cell including enzymes, membrane transport pumps, and ion channels. CaM interactions are involved in a multitude of cellular processes including, but not limited to, gene regulation, DNA synthesis, cell cycle  
20 progression, mitosis, cytokinesis, cytoskeletal organization, muscle contraction, signal transduction, ion homeostasis, exocytosis, and metabolic regulation (Celio, M.R. et al. (1996) Guidebook to Calcium-binding Proteins, Oxford University Press, Oxford, UK, pp. 15-20). Some CBPs can serve as a storage depot for  $\text{Ca}^{2+}$  in an inactive state. Calsequestrin is one such CBP that is expressed in isoforms specific to cardiac muscle and skeletal muscle. It is suggested that calsequestrin binds  $\text{Ca}^{2+}$   
25 in a rapidly exchangeable state that is released during  $\text{Ca}^{2+}$ -signaling conditions (Celio, M.R. et al. (1996) Guidebook to Calcium-binding Proteins, Oxford University Press, New York NY, pp. 222-224).

#### Cyclins

Cell division is the fundamental process by which all living things grow and reproduce. In  
30 most organisms, the cell cycle consists of three principle steps; interphase, mitosis, and cytokinesis. Interphase, involves preparations for cell division, replication of the DNA and production of essential proteins. In mitosis, the nuclear material is divided and separates to opposite sides of the cell. Cytokinesis is the final division and fission of the cell cytoplasm to produce the daughter cells.

The entry and exit of a cell from mitosis is regulated by the synthesis and destruction of a  
35 family of activating proteins called cyclins. Cyclins act by binding to and activating a group of cyclin-dependent protein kinases (Cdks) which then phosphorylate and activate selected proteins

involved in the mitotic process. Several types of cyclins exist. (Ciechanover, A. (1994) Cell 79:13-21.) Two principle types are mitotic cyclin, or cyclin B, which controls entry of the cell into mitosis, and G1 cyclin, which controls events that drive the cell out of mitosis.

#### Signal Complex Scaffolding Proteins

5           Certain proteins in intracellular signaling pathways serve to link or cluster other proteins involved in the signaling cascade. A conserved protein domain called the PDZ domain has been identified in various membrane-associated signaling proteins. This domain has been implicated in receptor and ion channel clustering and in the targeting of multiprotein signaling complexes to specialized functional regions of the cytosolic face of the plasma membrane. (For a review of PDZ  
10 domain-containing proteins, see Ponting, C.P. et al. (1997) Bioessays 19:469-479.) A large proportion of PDZ domains are found in the eukaryotic MAGUK (membrane-associated guanylate kinase) protein family, members of which bind to the intracellular domains of receptors and channels. However, PDZ domains are also found in diverse membrane-localized proteins such as protein tyrosine phosphatases, serine/threonine kinases, G-protein cofactors, and synapse-associated proteins  
15 such as syntrophins and neuronal nitric oxide synthase (nNOS). Generally, about one to three PDZ domains are found in a given protein, although up to nine PDZ domains have been identified in a single protein.

#### **Membrane Transport Molecules**

20           The plasma membrane acts as a barrier to most molecules. Transport between the cytoplasm and the extracellular environment, and between the cytoplasm and lumenal spaces of cellular organelles requires specific transport proteins. Each transport protein carries a particular class of molecule, such as ions, sugars, or amino acids, and often is specific to a certain molecular species of the class. A variety of human inherited diseases are caused by a mutation in a transport protein. For  
25 example, cystinuria is an inherited disease that results from the inability to transport cystine, the disulfide-linked dimer of cysteine, from the urine into the blood. Accumulation of cystine in the urine leads to the formation of cystine stones in the kidneys.

          Transport proteins are multi-pass transmembrane proteins, which either actively transport molecules across the membrane or passively allow them to cross. Active transport involves  
30 directional pumping of a solute across the membrane, usually against an electrochemical gradient. Active transport is tightly coupled to a source of metabolic energy, such as ATP hydrolysis or an electrochemically favorable ion gradient. Passive transport involves the movement of a solute down its electrochemical gradient. Transport proteins can be further classified as either carrier proteins or channel proteins. Carrier proteins, which can function in active or passive transport, bind to a  
35 specific solute to be transported and undergo a conformational change which transfers the bound solute across the membrane. Channel proteins, which only function in passive transport, form

hydrophilic pores across the membrane. When the pores open, specific solutes, such as inorganic ions, pass through the membrane and down the electrochemical gradient of the solute.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with  
5 simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous  
10  $\text{Na}^+/\text{K}^+$  ATPase. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging techniques and for  
15 specific targeting of radioisotopes to the thyroid gland (Levy, O. et al. (1997) Proc. Natl. Acad. Sci. USA 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) J. Biol. Chem. 273:7501-7506).

Transporters play a major role in the regulation of pH, excretion of drugs, and the cellular  
20  $\text{K}^+/\text{Na}^+$  balance. Monocarboxylate anion transporters are proton-coupled symporters with a broad substrate specificity that includes L-lactate, pyruvate, and the ketone bodies acetate, acetoacetate, and beta-hydroxybutyrate. At least seven isoforms have been identified to date. The isoforms are predicted to have twelve transmembrane (TM) helical domains with a large intracellular loop between TM6 and TM7, and play a critical role in maintaining intracellular pH by removing the protons that  
25 are produced stoichiometrically with lactate during glycolysis. The best characterized  $\text{H}^+$ -monocarboxylate transporter is that of the erythrocyte membrane, which transports L-lactate and a wide range of other aliphatic monocarboxylates. Other cells possess  $\text{H}^+$ -linked monocarboxylate transporters with differing substrate and inhibitor selectivities. In particular, cardiac muscle and tumor cells have transporters that differ in their  $K_m$  values for certain substrates,  
30 including stereoselectivity for L- over D-lactate, and in their sensitivity to inhibitors. There are  $\text{Na}^+$ -monocarboxylate cotransporters on the luminal surface of intestinal and kidney epithelia, which allow the uptake of lactate, pyruvate, and ketone bodies in these tissues. In addition, there are specific and selective transporters for organic cations and organic anions in organs including the kidney, intestine and liver. Organic anion transporters are selective for hydrophobic, charged  
35 molecules with electron-attracting side groups. Organic cation transporters, such as the ammonium transporter, mediate the secretion of a variety of drugs and endogenous metabolites, and contribute to

the maintenance of intercellular pH. (Poole, R.C. and A.P. Halestrap (1993) *Am. J. Physiol.* 264:C761-C782; Price, N.T. et al. (1998) *Biochem. J.* 329:321-328; and Martinelle, K. and I. Haggstrom (1993) *J. Biotechnol.* 30: 339-350.)

The largest and most diverse family of transport proteins known is the ATP-binding cassette (ABC) transporters. As a family, ABC transporters can transport substances that differ markedly in chemical structure and size, ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC proteins consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) *Meth. Enzymol.* 292:131-163).

Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant fraction of fatty acid transport appears to occur via a high affinity, low capacity protein-mediated transport process. Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates uptake of long-chain fatty acids (Hui, T.Y. et al. (1998) *J. Biol. Chem.* 273:27420-27429).

#### Ion Channels

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form an ion-selective pore within the membrane. There are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical



impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient.

5 These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including  $\text{Na}^+\text{-K}^+$  ATPase,  $\text{Ca}^{2+}$ -ATPase, and  $\text{H}^+$ -ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of  $\text{Na}^+$  and  $\text{Ca}^{2+}$  are low and cytosolic concentration of  $\text{K}^+$  is high. The vacuolar (V) class of ion transporters includes  $\text{H}^+$  pumps on intracellular  
10 organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of  $\text{H}^+$  pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate ( $\text{P}_i$ ).

The resting potential of the cell is utilized in many processes involving carrier proteins and  
15 gated ion channels. Carrier proteins utilize the resting potential to transport molecules into and out of the cell. Amino acid and glucose transport into many cells is linked to sodium ion co-transport (symport) so that the movement of  $\text{Na}^+$  down an electrochemical gradient drives transport of the other molecule up a concentration gradient. Similarly, cardiac muscle links transfer of  $\text{Ca}^{2+}$  out of the cell with transport of  $\text{Na}^+$  into the cell (antiport).

20 Ion channels share common structural and mechanistic themes. The channel consists of four or five subunits or protein monomers that are arranged like a barrel in the plasma membrane. Each subunit typically consists of six potential transmembrane segments (S1, S2, S3, S4, S5, and S6). The center of the barrel forms a pore lined by  $\alpha$ -helices or  $\beta$ -strands. The side chains of the amino acid residues comprising the  $\alpha$ -helices or  $\beta$ -strands establish the charge (cation or anion) selectivity of the  
25 channel. The degree of selectivity, or what specific ions are allowed to pass through the channel, depends on the diameter of the narrowest part of the pore.

Gated ion channels control ion flow by regulating the opening and closing of pores. These channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open pores in response to mechanical stress, voltage-gated channels open pores in  
30 response to changes in membrane potential, and ligand-gated channels open pores in the presence of a specific ion, nucleotide, or neurotransmitter.

Voltage-gated  $\text{Na}^+$  and  $\text{K}^+$  channels are necessary for the function of electrically excitable cells, such as nerve and muscle cells. Action potentials, which lead to neurotransmitter release and muscle contraction, arise from large, transient changes in the permeability of the membrane to  $\text{Na}^+$   
35 and  $\text{K}^+$  ions. Depolarization of the membrane beyond the threshold level opens voltage-gated  $\text{Na}^+$  channels. Sodium ions flow into the cell, further depolarizing the membrane and opening more

voltage-gated  $\text{Na}^+$  channels, which propagates the depolarization down the length of the cell. Depolarization also opens voltage-gated potassium channels. Consequently, potassium ions flow outward, which leads to repolarization of the membrane. Voltage-gated channels utilize charged residues in the fourth transmembrane segment (S4) to sense voltage change. The open state lasts only  
5 about 1 millisecond, at which time the channel spontaneously converts into an inactive state that cannot be opened irrespective of the membrane potential. Inactivation is mediated by the channel's N-terminus, which acts as a plug that closes the pore. The transition from an inactive to a closed state requires a return to resting potential.

Voltage-gated  $\text{Na}^+$  channels are heterotrimeric complexes composed of a 260 kDa pore  
10 forming  $\alpha$  subunit that associates with two smaller auxiliary subunits,  $\beta 1$  and  $\beta 2$ . The  $\beta 2$  subunit is an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with  $\alpha$  and  $\beta 1$  subunits correlates with increased functional expression of the channel, a change in its gating properties, and an increase in whole cell capacitance due to an increase in membrane surface area. (Isom, L.L. et al. (1995) Cell 83:433-442.)

15 Voltage-gated  $\text{Ca}^{2+}$  channels are involved in presynaptic neurotransmitter release, and heart and skeletal muscle contraction. The voltage-gated  $\text{Ca}^{2+}$  channels from skeletal muscle (L-type) and brain (N-type) have been purified, and though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The  $\alpha_1$  subunit forms the membrane pore and voltage sensor, while the  $\alpha_2\delta$  and  $\beta$  subunits modulate the voltage-dependence,  
20 gating properties, and the current amplitude of the channel. These subunits are encoded by at least six  $\alpha_1$ , one  $\alpha_2\delta$ , and four  $\beta$  genes. A fourth subunit,  $\gamma$ , has been identified in skeletal muscle. (Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; and Jay, S.D. et al. (1990) Science 248:490-492.)

Chloride channels are necessary in endocrine secretion and in regulation of cytosolic and  
25 organelle pH. In secretory epithelial cells,  $\text{Cl}^-$  enters the cell across a basolateral membrane through an  $\text{Na}^+$ ,  $\text{K}^+/\text{Cl}^-$  cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of  $\text{Cl}^-$  from the apical surface, in response to hormonal stimulation, leads to flow of  $\text{Na}^+$  and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal  
30 genetic disorder in humans. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus", and devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) J. Exp. Biol. 172:245-266).

35 Many intracellular organelles contain  $\text{H}^+$ -ATPase pumps that generate transmembrane pH and electrochemical differences by moving protons from the cytosol to the organelle lumen. If the

membrane of the organelle is permeable to other ions, then the electrochemical gradient can be abrogated without affecting the pH differential. In fact, removal of the electrochemical barrier allows more  $H^+$  to be pumped across the membrane, increasing the pH differential.  $Cl^-$  is the sole counterion of  $H^+$  translocation in a number of organelles, including chromaffin granules, Golgi vesicles, lysosomes, and endosomes. Functions that require a low vacuolar pH include uptake of small molecules such as biogenic amines in chromaffin granules, processing of vacuolar constituents such as pro-hormones by proteolytic enzymes, and protein degradation in lysosomes (Al-Awqati, supra).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of  $Na^+$  and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as  $\gamma$ -aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential.

Ligand-gated channels can be regulated by intracellular second messengers. Calcium-activated  $K^+$  channels are gated by internal calcium ions. In nerve cells, an influx of calcium during depolarization opens  $K^+$  channels to modulate the magnitude of the action potential (Ishi, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656). Cyclic nucleotide-gated (CNG) channels are gated by cytosolic cyclic nucleotides. The best examples of these are the cAMP-gated  $Na^+$  channels involved in olfaction and the cGMP-gated cation channels involved in vision. Both systems involve ligand-mediated activation of a G-protein coupled receptor which then alters the level of cyclic nucleotide within the cell.

Ion channels are expressed in a number of tissues where they are implicated in a variety of processes. CNG channels, while abundantly expressed in photoreceptor and olfactory sensory cells, are also found in kidney, lung, pineal, retinal ganglion cells, testis, aorta, and brain. Calcium-activated  $K^+$  channels may be responsible for the vasodilatory effects of bradykinin in the kidney and for shunting excess  $K^+$  from brain capillary endothelial cells into the blood. They are also implicated in repolarizing granulocytes after agonist-stimulated depolarization (Ishi, supra). Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98).

#### Disease Correlation

The etiology of numerous human diseases and disorders can be attributed to defects in the transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g. cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small molecules across membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) *Exp. Nephrol.* 4:253-262; Talente, G.M. et al. (1994) *Ann. Intern. Med.* 120:218-226; and Chillon, M. et al. (1995) *New Engl. J. Med.* 332:1475-1480).

## 10 Protein Modification and Maintenance Molecules

The cellular processes regulating modification and maintenance of protein molecules coordinate their conformation, stabilization, and degradation. Each of these processes is mediated by key enzymes or proteins such as proteases, protease inhibitors, transferases, isomerases, and molecular chaperones.

### 15 Proteases

Proteases cleave proteins and peptides at the peptide bond that forms the backbone of the peptide and protein chain. Proteolytic processing is essential to cell growth, differentiation, remodeling, and homeostasis as well as inflammation and immune response. Typical protein half-lives range from hours to a few days, so that within all living cells, precursor proteins are being  
20 cleaved to their active form, signal sequences proteolytically removed from targeted proteins, and aged or defective proteins degraded by proteolysis. Proteases function in bacterial, parasitic, and viral invasion and replication within a host. Four principal categories of mammalian proteases have been identified based on active site structure, mechanism of action, and overall three-dimensional structure. (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford  
25 University Press, New York NY, pp. 1-5).

The serine proteases (SPs) have a serine residue, usually within a conserved sequence, in an active site composed of the serine, an aspartate, and a histidine residue. SPs include the digestive enzymes trypsin and chymotrypsin, components of the complement cascade and the blood-clotting cascade, and enzymes that control extracellular protein degradation. The main SP sub-families are  
30 trypases, which cleave after arginine or lysine; aspartases, which cleave after aspartate; chymases, which cleave after phenylalanine or leucine; metases, which cleavage after methionine; and serases which cleave after serine. Enterokinase, the initiator of intestinal digestion, is a serine protease found in the intestinal brush border, where it cleaves the acidic propeptide from trypsinogen to yield active trypsin (Kitamoto, Y. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:7588-7592).

35 Prolylcarboxypeptidase, a lysosomal serine peptidase that cleaves peptides such as angiotensin II and III and [des-Arg9] bradykinin, shares sequence homology with members of both the serine

carboxypeptidase and prolylendopeptidase families (Tan, F. et al. (1993) J. Biol. Chem. 268:16631-16638).

Cysteine proteases (CPs) have a cysteine as the major catalytic residue at an active site where catalysis proceeds via an intermediate thiol ester and is facilitated by adjacent histidine and aspartic acid residues. CPs are involved in diverse cellular processes ranging from the processing of precursor proteins to intracellular degradation. Mammalian CPs include lysosomal cathepsins and cytosolic calcium activated proteases, calpains. CPs are produced by monocytes, macrophages and other cells of the immune system which migrate to sites of inflammation and secrete molecules involved in tissue repair. Overabundance of these repair molecules plays a role in certain disorders. In autoimmune diseases such as rheumatoid arthritis, secretion of the cysteine peptidase cathepsin C degrades collagen, laminin, elastin and other structural proteins found in the extracellular matrix of bones.

Aspartic proteases are members of the cathepsin family of lysosomal proteases and include pepsin A, gastricsin, chymosin, renin, and cathepsins D and E. Aspartic proteases have a pair of aspartic acid residues in the active site, and are most active in the pH 2 - 3 range, in which one of the aspartate residues is ionized, the other un-ionized. Aspartic proteases include bacterial penicillopepsin, mammalian pepsin, renin, chymosin, and certain fungal proteases. Abnormal regulation and expression of cathepsins is evident in various inflammatory disease states. In cells isolated from inflamed synovia, the mRNA for stromelysin, cytokines, TIMP-1, cathepsin, gelatinase, and other molecules is preferentially expressed. Expression of cathepsins L and D is elevated in synovial tissues from patients with rheumatoid arthritis and osteoarthritis. Cathepsin L expression may also contribute to the influx of mononuclear cells which exacerbates the destruction of the rheumatoid synovium. (Keyszer, G.M. (1995) Arthritis Rheum. 38:976-984.) The increased expression and differential regulation of the cathepsins are linked to the metastatic potential of a variety of cancers and as such are of therapeutic and prognostic interest (Chambers, A.F. et al. (1993) Crit. Rev. Oncog. 4:95-114).

Metalloproteases have active sites that include two glutamic acid residues and one histidine residue that serve as binding sites for zinc. Carboxypeptidases A and B are the principal mammalian metalloproteases. Both are exoproteases of similar structure and active sites. Carboxypeptidase A, like chymotrypsin, prefers C-terminal aromatic and aliphatic side chains of hydrophobic nature, whereas carboxypeptidase B is directed toward basic arginine and lysine residues. Glycoprotease (GCP), or O-sialoglycoprotein endopeptidase, is a metallopeptidase which specifically cleaves O-sialoglycoproteins such as glycophorin A. Another metallopeptidase, placental leucine aminopeptidase (P-LAP) degrades several peptide hormones such as oxytocin and vasopressin, suggesting a role in maintaining homeostasis during pregnancy, and is expressed in several tissues (Rogi, T. et al. (1996) J. Biol. Chem. 271:56-61).

Ubiquitin proteases are associated with the ubiquitin conjugation system (UCS), a major pathway for the degradation of cellular proteins in eukaryotic cells and some bacteria. The UCS mediates the elimination of abnormal proteins and regulates the half-lives of important regulatory proteins that control cellular processes such as gene transcription and cell cycle progression. In the UCS pathway, proteins targeted for degradation are conjugated to a ubiquitin, a small heat stable protein. The ubiquitinated protein is then recognized and degraded by proteasome, a large, multisubunit proteolytic enzyme complex, and ubiquitin is released for reutilization by ubiquitin protease. The UCS is implicated in the degradation of mitotic cyclic kinases, oncoproteins, tumor suppressor genes such as p53, viral proteins, cell surface receptors associated with signal transduction, transcriptional regulators, and mutated or damaged proteins (Ciechanover, A. (1994) Cell 79:13-21). A murine proto-oncogene, Unp, encodes a nuclear ubiquitin protease whose overexpression leads to oncogenic transformation of NIH3T3 cells, and the human homolog of this gene is consistently elevated in small cell tumors and adenocarcinomas of the lung (Gray, D.A. (1995) Oncogene 10:2179-2183).

#### 15 Signal Peptidases

The mechanism for the translocation process into the endoplasmic reticulum (ER) involves the recognition of an N-terminal signal peptide on the elongating protein. The signal peptide directs the protein and attached ribosome to a receptor on the ER membrane. The polypeptide chain passes through a pore in the ER membrane into the lumen while the N-terminal signal peptide remains attached at the membrane surface. The process is completed when signal peptidase located inside the ER cleaves the signal peptide from the protein and releases the protein into the lumen.

#### Protease Inhibitors

Protease inhibitors and other regulators of protease activity control the activity and effects of proteases. Protease inhibitors have been shown to control pathogenesis in animal models of proteolytic disorders (Murphy, G. (1991) Agents Actions Suppl. 35:69-76). Low levels of the cystatins, low molecular weight inhibitors of the cysteine proteases, correlate with malignant progression of tumors. (Calkins, C. et al (1995) Biol. Biochem. Hoppe Seyler 376:71-80). Serpins are inhibitors of mammalian plasma serine proteases. Many serpins serve to regulate the blood clotting cascade and/or the complement cascade in mammals. Sp32 is a positive regulator of the mammalian acrosomal protease, acrosin, that binds the proenzyme, proacrosin, and thereby aides in packaging the enzyme into the acrosomal matrix (Baba, T. et al. (1994) J. Biol. Chem. 269:10133-10140). The Kunitz family of serine protease inhibitors are characterized by one or more "Kunitz domains" containing a series of cysteine residues that are regularly spaced over approximately 50 amino acid residues and form three intrachain disulfide bonds. Members of this family include aprotinin, tissue factor pathway inhibitor (TFPI-1 and TFPI-2), inter- $\alpha$ -trypsin inhibitor, and bikunin. (Marlor, C.W. et al. (1997) J. Biol. Chem. 272:12202-12208.) Members of this family are potent

inhibitors (in the nanomolar range) against serine proteases such as kallikrein and plasmin. Aprotinin has clinical utility in reduction of perioperative blood loss.

A major portion of all proteins synthesized in eukaryotic cells are synthesized on the cytosolic surface of the endoplasmic reticulum (ER). Before these immature proteins are distributed  
5 to other organelles in the cell or are secreted, they must be transported into the interior lumen of the ER where post-translational modifications are performed. These modifications include protein folding and the formation of disulfide bonds, and N-linked glycosylations.

#### Protein Isomerases

Protein folding in the ER is aided by two principal types of protein isomerases, protein  
10 disulfide isomerase (PDI), and peptidyl-prolyl isomerase (PPI). PDI catalyzes the oxidation of free sulfhydryl groups in cysteine residues to form intramolecular disulfide bonds in proteins. PPI, an enzyme that catalyzes the isomerization of certain proline imidic bonds in oligopeptides and proteins, is considered to govern one of the rate limiting steps in the folding of many proteins to their final functional conformation. The cyclophilins represent a major class of PPI that was originally  
15 identified as the major receptor for the immunosuppressive drug cyclosporin A (Handschumacher, R.E. et al. (1984) Science 226: 544-547).

#### Protein Glycosylation

The glycosylation of most soluble secreted and membrane-bound proteins by oligosaccharides linked to asparagine residues in proteins is also performed in the ER. This reaction  
20 is catalyzed by a membrane-bound enzyme, oligosaccharyl transferase. Although the exact purpose of this "N-linked" glycosylation is unknown, the presence of oligosaccharides tends to make a glycoprotein resistant to protease digestion. In addition, oligosaccharides attached to cell-surface proteins called selectins are known to function in cell-cell adhesion processes (Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing Co., New York NY, p.608). "O-linked"  
25 glycosylation of proteins also occurs in the ER by the addition of N-acetylgalactosamine to the hydroxyl group of a serine or threonine residue followed by the sequential addition of other sugar residues to the first. This process is catalysed by a series of glycosyltransferases each specific for a particular donor sugar nucleotide and acceptor molecule (Lodish, H. et al. (1995) Molecular Cell Biology, W.H. Freeman and Co., New York NY, pp.700-708). In many cases, both N- and O-linked  
30 oligosaccharides appear to be required for the secretion of proteins or the movement of plasma membrane glycoproteins to the cell surface.

An additional glycosylation mechanism operates in the ER specifically to target lysosomal enzymes to lysosomes and prevent their secretion. Lysosomal enzymes in the ER receive an N-linked oligosaccharide, like plasma membrane and secreted proteins, but are then phosphorylated on  
35 one or two mannose residues. The phosphorylation of mannose residues occurs in two steps, the first step being the addition of an N-acetylglucosamine phosphate residue by N-acetylglucosamine

phosphotransferase, and the second the removal of the N-acetylglucosamine group by phosphodiesterase. The phosphorylated mannose residue then targets the lysosomal enzyme to a mannose 6-phosphate receptor which transports it to a lysosome vesicle (Lodish, supra, pp. 708-711).

### Chaperones

5           Molecular chaperones are proteins that aid in the proper folding of immature proteins and refolding of improperly folded ones, the assembly of protein subunits, and in the transport of unfolded proteins across membranes. Chaperones are also called heat-shock proteins (hsp) because of their tendency to be expressed in dramatically increased amounts following brief exposure of cells to elevated temperatures. This latter property most likely reflects their need in the refolding of  
10 proteins that have become denatured by the high temperatures. Chaperones may be divided into several classes according to their location, function, and molecular weight, and include hsp60, TCP1, hsp70, hsp40 (also called DnaJ), and hsp90. For example, hsp90 binds to steroid hormone receptors, represses transcription in the absence of the ligand, and provides proper folding of the ligand-binding domain of the receptor in the presence of the hormone (Burston, S.G. and A.R. Clarke (1995) *Essays*  
15 *Biochem.* 29:125-136). Hsp60 and hsp70 chaperones aid in the transport and folding of newly synthesized proteins. Hsp70 acts early in protein folding, binding a newly synthesized protein before it leaves the ribosome and transporting the protein to the mitochondria or ER before releasing the folded protein. Hsp60, along with hsp10, binds misfolded proteins and gives them the opportunity to refold correctly. All chaperones share an affinity for hydrophobic patches on incompletely folded  
20 proteins and the ability to hydrolyze ATP. The energy of ATP hydrolysis is used to release the hsp-bound protein in its properly folded state (Alberts, supra, pp 214, 571-572).

### **Nucleic Acid Synthesis and Modification Molecules**

#### Polymerases

25           DNA and RNA replication are critical processes for cell replication and function. DNA and RNA replication are mediated by the enzymes DNA and RNA polymerase, respectively, by a "templating" process in which the nucleotide sequence of a DNA or RNA strand is copied by complementary base-pairing into a complementary nucleic acid sequence of either DNA or RNA. However, there are fundamental differences between the two processes.

30           DNA polymerase catalyzes the stepwise addition of a deoxyribonucleotide to the 3'-OH end of a polynucleotide strand (the primer strand) that is paired to a second (template) strand. The new DNA strand therefore grows in the 5' to 3' direction (Alberts, B. et al. (1994) The Molecular Biology of the Cell, Garland Publishing Inc., New York NY, pp. 251-254). The substrates for the polymerization reaction are the corresponding deoxynucleotide triphosphates which must base-pair  
35 with the correct nucleotide on the template strand in order to be recognized by the polymerase. Because DNA exists as a double-stranded helix, each of the two strands may serve as a template for



the formation of a new complementary strand. Each of the two daughter cells of the dividing cell therefore inherits a new DNA double helix containing one old and one new strand. Thus, DNA is said to be replicated "semiconservatively" by DNA polymerase. In addition to the synthesis of new DNA, DNA polymerase is also involved in the repair of damaged DNA as discussed below under

5 "Ligases."

In contrast to DNA polymerase, RNA polymerase uses a DNA template strand to "transcribe" DNA into RNA using ribonucleotide triphosphates as substrates. Like DNA polymerization, RNA polymerization proceeds in a 5' to 3' direction by addition of a ribonucleoside monophosphate to the 3'-OH end of a growing RNA chain. DNA transcription generates messenger RNAs (mRNA) that

10 carry information for protein synthesis, as well as the transfer, ribosomal, and other RNAs that have structural or catalytic functions. In eukaryotes, three discrete RNA polymerases synthesize the three different types of RNA (Alberts, *supra*, pp. 367-368). RNA polymerase I makes the large ribosomal RNAs, RNA polymerase II makes the mRNAs that will be translated into proteins, and RNA polymerase III makes a variety of small, stable RNAs, including 5S ribosomal RNA and the transfer

15 RNAs (tRNA). In all cases, RNA synthesis is initiated by binding of the RNA polymerase to a promoter region on the DNA and synthesis begins at a start site within the promoter. Synthesis is completed at a broad, general stop or termination region in the DNA where both the polymerase and the completed RNA chain are released.

Ligases

20 DNA repair is the process by which accidental base changes, such as those produced by oxidative damage, hydrolytic attack, or uncontrolled methylation of DNA are corrected before replication or transcription of the DNA can occur. Because of the efficiency of the DNA repair process, fewer than one in one thousand accidental base changes causes a mutation (Alberts, *supra*, pp. 245-249). The three steps common to most types of DNA repair are (1) excision of the damaged

25 or altered base or nucleotide by DNA nucleases, leaving a gap; (2) insertion of the correct nucleotide in this gap by DNA polymerase using the complementary strand as the template; and (3) sealing the break left between the inserted nucleotide(s) and the existing DNA strand by DNA ligase. In the last reaction, DNA ligase uses the energy from ATP hydrolysis to activate the 5' end of the broken phosphodiester bond before forming the new bond with the 3'-OH of the DNA strand. In Bloom's

30 syndrome, an inherited human disease, individuals are partially deficient in DNA ligation and consequently have an increased incidence of cancer (Alberts, *supra*, p. 247).

Nucleases

Nucleases comprise both enzymes that hydrolyze DNA (DNase) and RNA (RNase). They serve different purposes in nucleic acid metabolism. Nucleases hydrolyze the phosphodiester bonds

35 between adjacent nucleotides either at internal positions (endonucleases) or at the terminal 3' or 5' nucleotide positions (exonucleases). A DNA exonuclease activity in DNA polymerase, for example,

serves to remove improperly paired nucleotides attached to the 3'-OH end of the growing DNA strand by the polymerase and thereby serves a "proofreading" function. As mentioned above, DNA endonuclease activity is involved in the excision step of the DNA repair process.

RNases also serve a variety of functions. For example, RNase P is a ribonucleoprotein enzyme which cleaves the 5' end of pre-tRNAs as part of their maturation process. RNase H digests the RNA strand of an RNA/DNA hybrid. Such hybrids occur in cells invaded by retroviruses, and RNase H is an important enzyme in the retroviral replication cycle. Pancreatic RNase secreted by the pancreas into the intestine hydrolyzes RNA present in ingested foods. RNase activity in serum and cell extracts is elevated in a variety of cancers and infectious diseases (Schein, C.H. (1997) Nat. Biotechnol. 15:529-536). Regulation of RNase activity is being investigated as a means to control tumor angiogenesis, allergic reactions, viral infection and replication, and fungal infections.

#### Methylases

Methylation of specific nucleotides occurs in both DNA and RNA, and serves different functions in the two macromolecules. Methylation of cytosine residues to form 5-methyl cytosine in DNA occurs specifically at CG sequences which are base-paired with one another in the DNA double-helix. This pattern of methylation is passed from generation to generation during DNA replication by an enzyme called "maintenance methylase" that acts preferentially on those CG sequences that are base-paired with a CG sequence that is already methylated. Such methylation appears to distinguish active from inactive genes by preventing the binding of regulatory proteins that "turn on" the gene, but permit the binding of proteins that inactivate the gene (Alberts, *supra*, pp. 448-451). In RNA metabolism, "tRNA methylase" produces one of several nucleotide modifications in tRNA that affect the conformation and base-pairing of the molecule and facilitate the recognition of the appropriate mRNA codons by specific tRNAs. The primary methylation pattern is the dimethylation of guanine residues to form N,N-dimethyl guanine.

#### Helicases and Single-Stranded Binding Proteins

Helicases are enzymes that destabilize and unwind double helix structures in both DNA and RNA. Since DNA replication occurs more or less simultaneously on both strands, the two strands must first separate to generate a replication "fork" for DNA polymerase to act on. Two types of replication proteins contribute to this process, DNA helicases and single-stranded binding proteins. DNA helicases hydrolyze ATP and use the energy of hydrolysis to separate the DNA strands. Single-stranded binding proteins (SSBs) then bind to the exposed DNA strands without covering the bases, thereby temporarily stabilizing them for templating by the DNA polymerase (Alberts, *supra*, pp. 255-256).

RNA helicases also alter and regulate RNA conformation and secondary structure. Like the DNA helicases, RNA helicases utilize energy derived from ATP hydrolysis to destabilize and unwind RNA duplexes. The most well-characterized and ubiquitous family of RNA helicases is the DEAD-

box family, so named for the conserved B-type ATP-binding motif which is diagnostic of proteins in this family. Over 40 DEAD-box helicases have been identified in organisms as diverse as bacteria, insects, yeast, amphibians, mammals, and plants. DEAD-box helicases function in diverse processes such as translation initiation, splicing, ribosome assembly, and RNA editing, transport, and stability.

- 5 Some DEAD-box helicases play tissue- and stage-specific roles in spermatogenesis and embryogenesis. Overexpression of the DEAD-box 1 protein (DDX1) may play a role in the progression of neuroblastoma (Nb) and retinoblastoma (Rb) tumors (Godbout, R. et al. (1998) *J. Biol. Chem.* 273:21161-21168). These observations suggest that DDX1 may promote or enhance tumor progression by altering the normal secondary structure and expression levels of RNA in cancer cells.
- 10 Other DEAD-box helicases have been implicated either directly or indirectly in tumorigenesis (Discussed in Godbout, *supra*). For example, murine p68 is mutated in ultraviolet light-induced tumors, and human DDX6 is located at a chromosomal breakpoint associated with B-cell lymphoma. Similarly, a chimeric protein comprised of DDX10 and NUP98, a nucleoporin protein, may be involved in the pathogenesis of certain myeloid malignancies.

15 Topoisomerases

- Besides the need to separate DNA strands prior to replication, the two strands must be "unwound" from one another prior to their separation by DNA helicases. This function is performed by proteins known as DNA topoisomerases. DNA topoisomerase effectively acts as a reversible nuclease that hydrolyzes a phosphodiesterase bond in a DNA strand, permitting the two strands to
- 20 rotate freely about one another to remove the strain of the helix, and then rejoins the original phosphodiester bond between the two strands. Two types of DNA topoisomerase exist, types I and II. DNA Topoisomerase I causes a single-strand break in a DNA helix to allow the rotation of the two strands of the helix about the remaining phosphodiester bond in the opposite strand. DNA topoisomerase II causes a transient break in both strands of a DNA helix where two double helices
- 25 cross over one another. This type of topoisomerase can efficiently separate two interlocked DNA circles (Alberts, *supra*, pp.260-262). Type II topoisomerases are largely confined to proliferating cells in eukaryotes, such as cancer cells. For this reason they are targets for anticancer drugs. Topoisomerase II has been implicated in multi-drug resistance (MDR) as it appears to aid in the repair of DNA damage inflicted by DNA binding agents such as doxorubicin and vincristine.

30 Recombinases

- Genetic recombination is the process of rearranging DNA sequences within an organism's genome to provide genetic variation for the organism in response to changes in the environment. DNA recombination allows variation in the particular combination of genes present in an individual's genome, as well as the timing and level of expression of these genes (see Alberts, *supra*, pp. 263-
- 35 273). Two broad classes of genetic recombination are commonly recognized, general recombination and site-specific recombination. General recombination involves genetic exchange between any

homologous pair of DNA sequences usually located on two copies of the same chromosome. The process is aided by enzymes called recombinases that "nick" one strand of a DNA duplex more or less randomly and permit exchange with the complementary strand of another duplex. The process does not normally change the arrangement of genes on a chromosome. In site-specific  
5 recombination, the recombinase recognizes specific nucleotide sequences present in one or both of the recombining molecules. Base-pairing is not involved in this form of recombination and therefore does not require DNA homology between the recombining molecules. Unlike general recombination, this form of recombination can alter the relative positions of nucleotide sequences in chromosomes.

#### Splicing Factors

10 Various proteins are necessary for processing of transcribed RNAs in the nucleus. Pre-mRNA processing steps include capping at the 5' end with methylguanosine, polyadenylating the 3' end, and splicing to remove introns. The primary RNA transcript from DNA is a faithful copy of the gene containing both exon and intron sequences, and the latter sequences must be cut out of the RNA transcript to produce an mRNA that codes for a protein. This "splicing" of the mRNA sequence takes  
15 place in the nucleus with the aid of a large, multicomponent ribonucleoprotein complex known as a spliceosome. The spliceosomal complex is composed of five small nuclear ribonucleoprotein particles (snRNPs) designated U1, U2, U4, U5, and U6, and a number of additional proteins. Each snRNP contains a single species of snRNA and about ten proteins. The RNA components of some snRNPs recognize and base pair with intron consensus sequences. The protein components mediate  
20 spliceosome assembly and the splicing reaction. Autoantibodies to snRNP proteins are found in the blood of patients with systemic lupus erythematosus (Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, p. 863).

#### **Adhesion Molecules**

25 The surface of a cell is rich in transmembrane proteoglycans, glycoproteins, glycolipids, and receptors. These macromolecules mediate adhesion with other cells and with components of the extracellular matrix (ECM). The interaction of the cell with its surroundings profoundly influences cell shape, strength, flexibility, motility, and adhesion. These dynamic properties are intimately associated with signal transduction pathways controlling cell proliferation and differentiation, tissue  
30 construction, and embryonic development.

#### Cadherins

Cadherins comprise a family of calcium-dependent glycoproteins that function in mediating cell-cell adhesion in virtually all solid tissues of multicellular organisms. These proteins share multiple repeats of a cadherin-specific motif, and the repeats form the folding units of the cadherin  
35 extracellular domain. Cadherin molecules cooperate to form focal contacts, or adhesion plaques, between adjacent epithelial cells. The cadherin family includes the classical cadherins and

protocadherins. Classical cadherins include the E-cadherin, N-cadherin, and P-cadherin subfamilies. E-cadherin is present on many types of epithelial cells and is especially important for embryonic development. N-cadherin is present on nerve, muscle, and lens cells and is also critical for embryonic development. P-cadherin is present on cells of the placenta and epidermis. Recent studies report that protocadherins are involved in a variety of cell-cell interactions (Suzuki, S.T. (1996) J. Cell Sci. 109:2609-2611). The intracellular anchorage of cadherins is regulated by their dynamic association with catenins, a family of cytoplasmic signal transduction proteins associated with the actin cytoskeleton. The anchorage of cadherins to the actin cytoskeleton appears to be regulated by protein tyrosine phosphorylation, and the cadherins are the target of phosphorylation-induced junctional disassembly (Aberle, H. et al. (1996) J. Cell. Biochem. 61:514-523).

### Integrins

Integrins are ubiquitous transmembrane adhesion molecules that link the ECM to the internal cytoskeleton. Integrins are composed of two noncovalently associated transmembrane glycoprotein subunits called  $\alpha$  and  $\beta$ . Integrins function as receptors that play a role in signal transduction. For example, binding of integrin to its extracellular ligand may stimulate changes in intracellular calcium levels or protein kinase activity (Sjaastad, M.D. and W.J. Nelson (1997) BioEssays 19:47-55). At least ten cell surface receptors of the integrin family recognize the ECM component fibronectin, which is involved in many different biological processes including cell migration and embryogenesis (Johansson, S. et al. (1997) Front. Biosci. 2:D126-D146).

### Lectins

Lectins comprise a ubiquitous family of extracellular glycoproteins which bind cell surface carbohydrates specifically and reversibly, resulting in the agglutination of cells (reviewed in Drickamer, K. and M.E. Taylor (1993) Annu. Rev. Cell Biol. 9:237-264). This function is particularly important for activation of the immune response. Lectins mediate the agglutination and mitogenic stimulation of lymphocytes at sites of inflammation (Lasky, L.A. (1991) J. Cell. Biochem. 45:139-146; Palletta, E. et al. (1989) J. Immunol. 143:2850-2857).

Lectins are further classified into subfamilies based on carbohydrate-binding specificity and other criteria. The galectin subfamily, in particular, includes lectins that bind  $\beta$ -galactoside carbohydrate moieties in a thiol-dependent manner (reviewed in Hadari, Y.R. et al. (1998) J. Biol. Chem. 270:3447-3453). Galectins are widely expressed and developmentally regulated. Because all galectins lack an N-terminal signal peptide, it is suggested that galectins are externalized through an atypical secretory mechanism. Two classes of galectins have been defined based on molecular weight and oligomerization properties. Small galectins form homodimers and are about 14 to 16 kilodaltons in mass, while large galectins are monomeric and about 29-37 kilodaltons.

Galectins contain a characteristic carbohydrate recognition domain (CRD). The CRD is about 140 amino acids and contains several stretches of about 1 - 10 amino acids which are highly

conserved among all galectins. A particular 6-amino acid motif within the CRD contains conserved tryptophan and arginine residues which are critical for carbohydrate binding. The CRD of some galectins also contains cysteine residues which may be important for disulfide bond formation. Secondary structure predictions indicate that the CRD forms several  $\beta$ -sheets.

5        Galectins play a number of roles in diseases and conditions associated with cell-cell and cell-matrix interactions. For example, certain galectins associate with sites of inflammation and bind to cell surface immunoglobulin E molecules. In addition, galectins may play an important role in cancer metastasis. Galectin overexpression is correlated with the metastatic potential of cancers in humans and mice. Moreover, anti-galectin antibodies inhibit processes associated with cell  
10 transformation, such as cell aggregation and anchorage-independent growth (See, for example, Su, Z.-Z. et al. (1996) Proc. Natl. Acad. Sci. USA 93:7252-7257).

#### Selectins

      Selectins, or LEC-CAMs, comprise a specialized lectin subfamily involved primarily in inflammation and leukocyte adhesion (Reviewed in Lasky, *supra*). Selectins mediate the recruitment  
15 of leukocytes from the circulation to sites of acute inflammation and are expressed on the surface of vascular endothelial cells in response to cytokine signaling. Selectins bind to specific ligands on the leukocyte cell membrane and enable the leukocyte to adhere to and migrate along the endothelial surface. Binding of selectin to its ligand leads to polarized rearrangement of the actin cytoskeleton and stimulates signal transduction within the leukocyte (Brenner, B. et al. (1997) Biochem. Biophys.  
20 Res. Commun. 231:802-807; Hidari, K.I. et al. (1997) J. Biol. Chem. 272:28750-28756). Members of the selectin family possess three characteristic motifs: a lectin or carbohydrate recognition domain; an epidermal growth factor-like domain; and a variable number of short consensus repeats (scr or "sushi" repeats) which are also present in complement regulatory proteins. The selectins include lymphocyte adhesion molecule-1 (Lam-1 or L-selectin), endothelial leukocyte adhesion  
25 molecule-1 (ELAM-1 or E-selectin), and granule membrane protein-140 (GMP-140 or P-selectin) (Johnston, G.I. et al. (1989) Cell 56:1033-1044).

#### **Antigen Recognition Molecules**

      All vertebrates have developed sophisticated and complex immune systems that provide  
30 protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved  
35 structural domain. This Ig domain is comprised of antiparallel  $\beta$  sheets joined by a disulfide bond in an arrangement called the Ig fold. Members of the Ig superfamily include T-cell receptors, major

histocompatibility (MHC) proteins, antibodies, and immune cell-specific surface markers such as CD4, CD8, and CD28.

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York NY, pp. 1229-1246.)

Antibodies, or immunoglobulins, are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ , and  $\mu$  H-chain types. There are two types of L-chains,  $\kappa$  and  $\lambda$ , either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, *supra*, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain repeated Ig domains. For example, a typical H-chain

contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region.

5       The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene  
10 segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

T-cell receptors are both structurally and functionally related to antibodies. (Reviewed in  
15   Alberts, *supra*, pp. 1228-1229.) T-cell receptors are cell surface proteins that bind foreign antigens and mediate diverse aspects of the immune response. A typical T-cell receptor is a heterodimer comprised of two disulfide-linked polypeptide chains called  $\alpha$  and  $\beta$ . Each chain is about 280 amino acids in length and contains one variable region and one constant region. Each variable or constant region folds into an Ig domain. The variable regions from the  $\alpha$  and  $\beta$  chains come together in the  
20 heterodimer to form the antigen recognition site. T-cell receptor diversity is generated by somatic rearrangement of gene segments encoding the  $\alpha$  and  $\beta$  chains. T-cell receptors recognize small peptide antigens that are expressed on the surface of antigen-presenting cells and pathogen-infected cells. These peptide antigens are presented on the cell surface in association with major histocompatibility proteins which provide the proper context for antigen recognition.

25

#### **Secreted and Extracellular Matrix Molecules**

Protein secretion is essential for cellular function. Protein secretion is mediated by a signal peptide located at the amino terminus of the protein to be secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to  
30 the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Secreted proteins are often synthesized as inactive precursors that are activated by post-translational processing events during transit through  
35 the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include



chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides include receptors, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, neuropeptides, vasomediators, ion channels, transporters/pumps, and proteases. (Reviewed in

5   Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York NY, pp. 557-560, 582-592.)

The extracellular matrix (ECM) is a complex network of glycoproteins, polysaccharides, proteoglycans, and other macromolecules that are secreted from the cell into the extracellular space. The ECM remains in close association with the cell surface and provides a supportive meshwork that

10   profoundly influences cell shape, motility, strength, flexibility, and adhesion. In fact, adhesion of a cell to its surrounding matrix is required for cell survival except in the case of metastatic tumor cells, which have overcome the need for cell-ECM anchorage. This phenomenon suggests that the ECM plays a critical role in the molecular mechanisms of growth control and metastasis. (Reviewed in Ruoslahti, E. (1996) *Sci. Am.* 275:72-77.) Furthermore, the ECM determines the structure and

15   physical properties of connective tissue and is particularly important for morphogenesis and other processes associated with embryonic development and pattern formation.

The collagens comprise a family of ECM proteins that provide structure to bone, teeth, skin, ligaments, tendons, cartilage, blood vessels, and basement membranes. Multiple collagen proteins have been identified. Three collagen molecules fold together in a triple helix stabilized by interchain

20   disulfide bonds. Bundles of these triple helices then associate to form fibrils. Collagen primary structure consists of hundreds of (Gly-X-Y) repeats where about a third of the X and Y residues are Pro. Glycines are crucial to helix formation as the bulkier amino acid sidechains cannot fold into the triple helical conformation. Because of these strict sequence requirements, mutations in collagen genes have severe consequences. Osteogenesis imperfecta patients have brittle bones that fracture

25   easily; in severe cases patients die in utero or at birth. Ehlers-Danlos syndrome patients have hyperelastic skin, hypermobile joints, and susceptibility to aortic and intestinal rupture. Chondrodysplasia patients have short stature and ocular disorders. Alport syndrome patients have hematuria, sensorineural deafness, and eye lens deformation. (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, Inc., New York NY, pp. 2105-2117; and

30   Creighton, T.E. (1984) Proteins, Structures and Molecular Principles, W.H. Freeman and Company, New York NY, pp. 191-197.)

Elastin and related proteins confer elasticity to tissues such as skin, blood vessels, and lungs. Elastin is a highly hydrophobic protein of about 750 amino acids that is rich in proline and glycine residues. Elastin molecules are highly cross-linked, forming an extensive extracellular network of

35   fibers and sheets. Elastin fibers are surrounded by a sheath of microfibrils which are composed of a number of glycoproteins, including fibrillin. Mutations in the gene encoding fibrillin are responsible

for Marfan's syndrome, a genetic disorder characterized by defects in connective tissue. In severe cases, the aortas of afflicted individuals are prone to rupture. (Reviewed in Alberts, *supra*, pp. 984-986.)

Fibronectin is a large ECM glycoprotein found in all vertebrates. Fibronectin exists as a dimer of two subunits, each containing about 2,500 amino acids. Each subunit folds into a rod-like structure containing multiple domains. The domains each contain multiple repeated modules, the most common of which is the type III fibronectin repeat. The type III fibronectin repeat is about 90 amino acids in length and is also found in other ECM proteins and in some plasma membrane and cytoplasmic proteins. Furthermore, some type III fibronectin repeats contain a characteristic tripeptide consisting of Arginine-Glycine-Aspartic acid (RGD). The RGD sequence is recognized by the integrin family of cell surface receptors and is also found in other ECM proteins. Disruption of both copies of the gene encoding fibronectin causes early embryonic lethality in mice. The mutant embryos display extensive morphological defects, including defects in the formation of the notochord, somites, heart, blood vessels, neural tube, and extraembryonic structures. (Reviewed in Alberts, *supra*, pp. 986-987.)

Laminin is a major glycoprotein component of the basal lamina which underlies and supports epithelial cell sheets. Laminin is one of the first ECM proteins synthesized in the developing embryo. Laminin is an 850 kilodalton protein composed of three polypeptide chains joined in the shape of a cross by disulfide bonds. Laminin is especially important for angiogenesis and in particular, for guiding the formation of capillaries. (Reviewed in Alberts, *supra*, pp. 990-991.)

There are many other types of proteinaceous ECM components, most of which can be classified as proteoglycans. Proteoglycans are composed of unbranched polysaccharide chains (glycosaminoglycans) attached to protein cores. Common proteoglycans include aggrecan, betaglycan, decorin, perlecan, serglycin, and syndecan-1. Some of these molecules not only provide mechanical support, but also bind to extracellular signaling molecules, such as fibroblast growth factor and transforming growth factor  $\beta$ , suggesting a role for proteoglycans in cell-cell communication and cell growth. (Reviewed in Alberts, *supra*, pp. 973-978.) Likewise, the glycoproteins tenascin-C and tenascin-R are expressed in developing and lesioned neural tissue and provide stimulatory and anti-adhesive (inhibitory) properties, respectively, for axonal growth. (Faissner, A. (1997) Cell Tissue Res. 290:331-341.)

### Cytoskeletal Molecules

The cytoskeleton is a cytoplasmic network of protein fibers that mediate cell shape, structure, and movement. The cytoskeleton supports the cell membrane and forms tracks along which organelles and other elements move in the cytosol. The cytoskeleton is a dynamic structure that allows cells to adopt various shapes and to carry out directed movements. Major cytoskeletal fibers

include the microtubules, the microfilaments, and the intermediate filaments. Motor proteins, including myosin, dynein, and kinesin, drive movement of or along the fibers. The motor protein dynamin drives the formation of membrane vesicles. Accessory or associated proteins modify the structure or activity of the fibers while cytoskeletal membrane anchors connect the fibers to the cell membrane.

### Tubulins

Microtubules, cytoskeletal fibers with a diameter of about 24 nm, have multiple roles in the cell. Bundles of microtubules form cilia and flagella, which are whip-like extensions of the cell membrane that are necessary for sweeping materials across an epithelium and for swimming of sperm, respectively. Marginal bands of microtubules in red blood cells and platelets are important for these cells' pliability. Organelles, membrane vesicles, and proteins are transported in the cell along tracks of microtubules. For example, microtubules run through nerve cell axons, allowing bi-directional transport of materials and membrane vesicles between the cell body and the nerve terminal. Failure to supply the nerve terminal with these vesicles blocks the transmission of neural signals. Microtubules are also critical to chromosomal movement during cell division. Both stable and short-lived populations of microtubules exist in the cell.

Microtubules are polymers of GTP-binding tubulin protein subunits. Each subunit is a heterodimer of  $\alpha$ - and  $\beta$ - tubulin, multiple isoforms of which exist. The hydrolysis of GTP is linked to the addition of tubulin subunits at the end of a microtubule. The subunits interact head to tail to form protofilaments; the protofilaments interact side to side to form a microtubule. A microtubule is polarized, one end ringed with  $\alpha$ -tubulin and the other with  $\beta$ -tubulin, and the two ends differ in their rates of assembly. Generally, each microtubule is composed of 13 protofilaments although 11 or 15 protofilament-microtubules are sometimes found. Cilia and flagella contain doublet microtubules. Microtubules grow from specialized structures known as centrosomes or microtubule-organizing centers (MTOCs). MTOCs may contain one or two centrioles, which are pinwheel arrays of triplet microtubules. The basal body, the organizing center located at the base of a cilium or flagellum, contains one centriole. Gamma tubulin present in the MTOC is important for nucleating the polymerization of  $\alpha$ - and  $\beta$ - tubulin heterodimers but does not polymerize into microtubules.

### Microtubule-Associated Proteins

Microtubule-associated proteins (MAPs) have roles in the assembly and stabilization of microtubules. One major family of MAPs, assembly MAPs, can be identified in neurons as well as non-neuronal cells. Assembly MAPs are responsible for cross-linking microtubules in the cytosol. These MAPs are organized into two domains: a basic microtubule-binding domain and an acidic projection domain. The projection domain is the binding site for membranes, intermediate filaments, or other microtubules. Based on sequence analysis, assembly MAPs can be further grouped into two types: Type I and Type II. Type I MAPs, which include MAP1A and MAP1B, are large, filamentous

molecules that co-purify with microtubules and are abundantly expressed in brain and testes. Type I MAPs contain several repeats of a positively-charged amino acid sequence motif that binds and neutralizes negatively charged tubulin, leading to stabilization of microtubules. MAP1A and MAP1B are each derived from a single precursor polypeptide that is subsequently proteolytically processed to generate one heavy chain and one light chain.

Another light chain, LC3, is a 16.4 kDa molecule that binds MAP1A, MAP1B, and microtubules. It is suggested that LC3 is synthesized from a source other than the MAP1A or MAP1B transcripts, and that the expression of LC3 may be important in regulating the microtubule binding activity of MAP1A and MAP1B during cell proliferation (Mann, S.S. et al. (1994) J. Biol. Chem. 269:11492-11497).

Type II MAPs, which include MAP2a, MAP2b, MAP2c, MAP4, and Tau, are characterized by three to four copies of an 18-residue sequence in the microtubule-binding domain. MAP2a, MAP2b, and MAP2c are found only in dendrites, MAP4 is found in non-neuronal cells, and Tau is found in axons and dendrites of nerve cells. Alternative splicing of the Tau mRNA leads to the existence of multiple forms of Tau protein. Tau phosphorylation is altered in neurodegenerative disorders such as Alzheimer's disease, Pick's disease, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia and Parkinsonism linked to chromosome 17. The altered Tau phosphorylation leads to a collapse of the microtubule network and the formation of intraneuronal Tau aggregates (Spillantini, M.G. and M. Goedert (1998) Trends Neurosci. 21:428-433).

The protein pericentrin is found in the MTOC and has a role in microtubule assembly.

#### Actins

Microfilaments, cytoskeletal filaments with a diameter of about 7-9 nm, are vital to cell locomotion, cell shape, cell adhesion, cell division, and muscle contraction. Assembly and disassembly of the microfilaments allow cells to change their morphology. Microfilaments are the polymerized form of actin, the most abundant intracellular protein in the eukaryotic cell. Human cells contain six isoforms of actin. The three  $\alpha$ -actins are found in different kinds of muscle, nonmuscle  $\beta$ -actin and nonmuscle  $\gamma$ -actin are found in nonmuscle cells, and another  $\gamma$ -actin is found in intestinal smooth muscle cells. G-actin, the monomeric form of actin, polymerizes into polarized, helical F-actin filaments, accompanied by the hydrolysis of ATP to ADP. Actin filaments associate to form bundles and networks, providing a framework to support the plasma membrane and determine cell shape. These bundles and networks are connected to the cell membrane. In muscle cells, thin filaments containing actin slide past thick filaments containing the motor protein myosin during contraction. A family of actin-related proteins exist that are not part of the actin cytoskeleton, but rather associate with microtubules and dynein.

#### Actin-Associated Proteins

Actin-associated proteins have roles in cross-linking, severing, and stabilization of actin filaments and in sequestering actin monomers. Several of the actin-associated proteins have multiple functions. Bundles and networks of actin filaments are held together by actin cross-linking proteins. These proteins have two actin-binding sites, one for each filament. Short cross-linking proteins  
5 promote bundle formation while longer, more flexible cross-linking proteins promote network formation. Calmodulin-like calcium-binding domains in actin cross-linking proteins allow calcium regulation of cross-linking. Group I cross-linking proteins have unique actin-binding domains and include the 30 kD protein, EF-1a, fascin, and scruin. Group II cross-linking proteins have a 7,000-MW actin-binding domain and include villin and dematin. Group III cross-linking proteins have  
10 pairs of a 26,000-MW actin-binding domain and include fimbrin, spectrin, dystrophin, ABP 120, and filamin.

Severing proteins regulate the length of actin filaments by breaking them into short pieces or by blocking their ends. Severing proteins include gCAP39, severin (fragmin), gelsolin, and villin. Capping proteins can cap the ends of actin filaments, but cannot break filaments. Capping proteins  
15 include CapZ and tropomodulin. The proteins thymosin and profilin sequester actin monomers in the cytosol, allowing a pool of unpolymerized actin to exist. The actin-associated proteins tropomyosin, troponin, and caldesmon regulate muscle contraction in response to calcium.

#### Intermediate Filaments and Associated Proteins

Intermediate filaments (IFs) are cytoskeletal fibers with a diameter of about 10 nm,  
20 intermediate between that of microfilaments and microtubules. IFs serve structural roles in the cell, reinforcing cells and organizing cells into tissues. IFs are particularly abundant in epidermal cells and in neurons. IFs are extremely stable, and, in contrast to microfilaments and microtubules, do not function in cell motility.

Five types of IF proteins are known in mammals. Type I and Type II proteins are the acidic  
25 and basic keratins, respectively. Heterodimers of the acidic and basic keratins are the building blocks of keratin IFs. Keratins are abundant in soft epithelia such as skin and cornea, hard epithelia such as nails and hair, and in epithelia that line internal body cavities. Mutations in keratin genes lead to epithelial diseases including epidermolysis bullosa simplex, bullous congenital ichthyosiform erythroderma (epidermolytic hyperkeratosis), non-epidermolytic and epidermolytic palmoplantar  
30 keratoderma, ichthyosis bullosa of Siemens, pachyonychia congenita, and white sponge nevus. Some of these diseases result in severe skin blistering. (See, e.g., Wawersik, M. et al. (1997) J. Biol. Chem. 272:32557-32565; and Corden L.D. and W.H. McLean (1996) Exp. Dermatol. 5:297-307.)

Type III IF proteins include desmin, glial fibrillary acidic protein, vimentin, and peripherin. Desmin filaments in muscle cells link myofibrils into bundles and stabilize sarcomeres in contracting  
35 muscle. Glial fibrillary acidic protein filaments are found in the glial cells that surround neurons and astrocytes. Vimentin filaments are found in blood vessel endothelial cells, some epithelial cells, and

mesenchymal cells such as fibroblasts, and are commonly associated with microtubules. Vimentin filaments may have roles in keeping the nucleus and other organelles in place in the cell. Type IV IFs include the neurofilaments and nestin. Neurofilaments, composed of three polypeptides NF-L, NF-M, and NF-H, are frequently associated with microtubules in axons. Neurofilaments are responsible for the radial growth and diameter of an axon, and ultimately for the speed of nerve impulse transmission. Changes in phosphorylation and metabolism of neurofilaments are observed in neurodegenerative diseases including amyotrophic lateral sclerosis, Parkinson's disease, and Alzheimer's disease (Julien, J.P. and W.E. Mushynski (1998) Prog. Nucleic Acid Res. Mol. Biol. 61:1-23). Type V IFs, the lamins, are found in the nucleus where they support the nuclear membrane.

IFs have a central  $\alpha$ -helical rod region interrupted by short nonhelical linker segments. The rod region is bracketed, in most cases, by non-helical head and tail domains. The rod regions of intermediate filament proteins associate to form a coiled-coil dimer. A highly ordered assembly process leads from the dimers to the IFs. Neither ATP nor GTP is needed for IF assembly, unlike that of microfilaments and microtubules.

IF-associated proteins (IFAPs) mediate the interactions of IFs with one another and with other cell structures. IFAPs cross-link IFs into a bundle, into a network, or to the plasma membrane, and may cross-link IFs to the microfilament and microtubule cytoskeleton. Microtubules and IFs are in particular closely associated. IFAPs include BPAG1, plakoglobin, desmoplakin I, desmoplakin II, plectin, ankyrin, filaggrin, and lamin B receptor.

#### Cytoskeletal-Membrane Anchors

Cytoskeletal fibers are attached to the plasma membrane by specific proteins. These attachments are important for maintaining cell shape and for muscle contraction. In erythrocytes, the spectrin-actin cytoskeleton is attached to cell membrane by three proteins, band 4.1, ankyrin, and adducin. Defects in this attachment result in abnormally shaped cells which are more rapidly degraded by the spleen, leading to anemia. In platelets, the spectrin-actin cytoskeleton is also linked to the membrane by ankyrin; a second actin network is anchored to the membrane by filamin. In muscle cells the protein dystrophin links actin filaments to the plasma membrane; mutations in the dystrophin gene lead to Duchenne muscular dystrophy. In adherens junctions and adhesion plaques the peripheral membrane proteins  $\alpha$ -actinin and vinculin attach actin filaments to the cell membrane.

IFs are also attached to membranes by cytoskeletal-membrane anchors. The nuclear lamina is attached to the inner surface of the nuclear membrane by the lamin B receptor. Vimentin IFs are attached to the plasma membrane by ankyrin and plectin. Desmosome and hemidesmosome membrane junctions hold together epithelial cells of organs and skin. These membrane junctions allow shear forces to be distributed across the entire epithelial cell layer, thus providing strength and rigidity to the epithelium. IFs in epithelial cells are attached to the desmosome by plakoglobin and desmoplakins. The proteins that link IFs to hemidesmosomes are not known. Desmin IFs surround

the sarcomere in muscle and are linked to the plasma membrane by paranemin, synemin, and ankyrin.

#### Myosin-related Motor Proteins

Myosins are actin-activated ATPases, found in eukaryotic cells, that couple hydrolysis of ATP with motion. Myosin provides the motor function for muscle contraction and intracellular movements such as phagocytosis and rearrangement of cell contents during mitotic cell division (cytokinesis). The contractile unit of skeletal muscle, termed the sarcomere, consists of highly ordered arrays of thin actin-containing filaments and thick myosin-containing filaments. Crossbridges form between the thick and thin filaments, and the ATP-dependent movement of myosin heads within the thick filaments pulls the thin filaments, shortening the sarcomere and thus the muscle fiber.

Myosins are composed of one or two heavy chains and associated light chains. Myosin heavy chains contain an amino-terminal motor or head domain, a neck that is the site of light-chain binding, and a carboxy-terminal tail domain. The tail domains may associate to form an  $\alpha$ -helical coiled coil. Conventional myosins, such as those found in muscle tissue, are composed of two myosin heavy-chain subunits, each associated with two light-chain subunits that bind at the neck region and play a regulatory role. Unconventional myosins, believed to function in intracellular motion, may contain either one or two heavy chains and associated light chains. There is evidence for about 25 myosin heavy chain genes in vertebrates, more than half of them unconventional.

#### Dynein-related Motor Proteins

Dyneins are (-) end-directed motor proteins which act on microtubules. Two classes of dyneins, cytosolic and axonemal, have been identified. Cytosolic dyneins are responsible for translocation of materials along cytoplasmic microtubules, for example, transport from the nerve terminal to the cell body and transport of endocytic vesicles to lysosomes. Cytoplasmic dyneins are also reported to play a role in mitosis. Axonemal dyneins are responsible for the beating of flagella and cilia. Dynein on one microtubule doublet walks along the adjacent microtubule doublet. This sliding force produces bending forces that cause the flagellum or cilium to beat. Dyneins have a native mass between 1000 and 2000 kDa and contain either two or three force-producing heads driven by the hydrolysis of ATP. The heads are linked via stalks to a basal domain which is composed of a highly variable number of accessory intermediate and light chains.

#### Kinesin-related Motor Proteins

Kinesins are (+) end-directed motor proteins which act on microtubules. The prototypical kinesin molecule is involved in the transport of membrane-bound vesicles and organelles. This function is particularly important for axonal transport in neurons. Kinesin is also important in all cell types for the transport of vesicles from the Golgi complex to the endoplasmic reticulum. This role is critical for maintaining the identity and functionality of these secretory organelles.

Kinesins define a ubiquitous, conserved family of over 50 proteins that can be classified into

at least 8 subfamilies based on primary amino acid sequence, domain structure, velocity of movement, and cellular function. (Reviewed in Moore, J.D. and S.A. Endow (1996) Bioessays 18:207-219; and Hoyt, A.M. (1994) Curr. Opin. Cell Biol. 6:63-68.) The prototypical kinesin molecule is a heterotetramer comprised of two heavy polypeptide chains (KHCs) and two light polypeptide chains (KLCs). The KHC subunits are typically referred to as "kinesin." KHC is about 1000 amino acids in length, and KLC is about 550 amino acids in length. Two KHCs dimerize to form a rod-shaped molecule with three distinct regions of secondary structure. At one end of the molecule is a globular motor domain that functions in ATP hydrolysis and microtubule binding. Kinesin motor domains are highly conserved and share over 70% identity. Beyond the motor domain is an  $\alpha$ -helical coiled-coil region which mediates dimerization. At the other end of the molecule is a fan-shaped tail that associates with molecular cargo. The tail is formed by the interaction of the KHC C-termini with the two KLCs.

Members of the more divergent subfamilies of kinesins are called kinesin-related proteins (KRPs), many of which function during mitosis in eukaryotes (Hoyt, *supra*). Some KRPs are required for assembly of the mitotic spindle. *In vivo* and *in vitro* analyses suggest that these KRPs exert force on microtubules that comprise the mitotic spindle, resulting in the separation of spindle poles. Phosphorylation of KRP is required for this activity. Failure to assemble the mitotic spindle results in abortive mitosis and chromosomal aneuploidy, the latter condition being characteristic of cancer cells. In addition, a unique KRP, centromere protein E, localizes to the kinetochore of human mitotic chromosomes and may play a role in their segregation to opposite spindle poles.

#### Dynamin-related Motor Proteins

Dynamin is a large GTPase motor protein that functions as a "molecular pinchase," generating a mechanochemical force used to sever membranes. This activity is important in forming clathrin-coated vesicles from coated pits in endocytosis and in the biogenesis of synaptic vesicles in neurons. Binding of dynamin to a membrane leads to dynamin's self-assembly into spirals that may act to constrict a flat membrane surface into a tubule. GTP hydrolysis induces a change in conformation of the dynamin polymer that pinches the membrane tubule, leading to severing of the membrane tubule and formation of a membrane vesicle. Release of GDP and inorganic phosphate leads to dynamin disassembly. Following disassembly the dynamin may either dissociate from the membrane or remain associated to the vesicle and be transported to another region of the cell. Three homologous dynamin genes have been discovered, in addition to several dynamin-related proteins. Conserved dynamin regions are the N-terminal GTP-binding domain, a central pleckstrin homology domain that binds membranes, a central coiled-coil region that may activate dynamin's GTPase activity, and a C-terminal proline-rich domain that contains several motifs that bind SH3 domains on other proteins. Some dynamin-related proteins do not contain the pleckstrin homology domain or the proline-rich domain. (See McNiven, M.A. (1998) Cell 94:151-154; Scaife, R.M. and R.L. Margolis



(1997) Cell. Signal. 9:395-401.)

The cytoskeleton is reviewed in Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY.

## 5 Ribosomal Molecules

Ribosomal RNAs (rRNAs) are assembled, along with ribosomal proteins, into ribosomes, which are cytoplasmic particles that translate messenger RNA into polypeptides. The eukaryotic ribosome is composed of a 60S (large) subunit and a 40S (small) subunit, which together form the 80S ribosome. In addition to the 18S, 28S, 5S, and 5.8S rRNAs, the ribosome also contains more  
10 than fifty proteins. The ribosomal proteins have a prefix which denotes the subunit to which they belong, either L (large) or S (small). Ribosomal protein activities include binding rRNA and organizing the conformation of the junctions between rRNA helices (Woodson, S.A. and N.B. Leontis (1998) Curr. Opin. Struct. Biol. 8:294-300; Ramakrishnan, V. and S.W. White (1998) Trends Biochem. Sci. 23:208-212.) Three important sites are identified on the ribosome. The aminoacyl-  
15 tRNA site (A site) is where charged tRNAs (with the exception of the initiator-tRNA) bind on arrival at the ribosome. The peptidyl-tRNA site (P site) is where new peptide bonds are formed, as well as where the initiator tRNA binds. The exit site (E site) is where deacylated tRNAs bind prior to their release from the ribosome. (The ribosome is reviewed in Stryer, L. (1995) Biochemistry W.H. Freeman and Company, New York NY, pp. 888-908; and Lodish, H. et al. (1995) Molecular Cell  
20 Biology Scientific American Books, New York NY. pp. 119-138.)

## Chromatin Molecules

The nuclear DNA of eukaryotes is organized into chromatin. Two types of chromatin are observed: euchromatin, some of which may be transcribed, and heterochromatin so densely packed  
25 that much of it is inaccessible to transcription. Chromatin packing thus serves to regulate protein expression in eukaryotes. Bacteria lack chromatin and the chromatin-packing level of gene regulation.

The fundamental unit of chromatin is the nucleosome of 200 DNA base pairs associated with two copies each of histones H2A, H2B, H3, and H4. Adjacent nucleosomes are linked by another  
30 class of histones, H1. Low molecular weight non-histone proteins called the high mobility group (HMG), associated with chromatin, may function in the unwinding of DNA and stabilization of single-stranded DNA. Chromodomain proteins function in compaction of chromatin into its transcriptionally silent heterochromatin form.

During mitosis, all DNA is compacted into heterochromatin and transcription ceases.  
35 Transcription in interphase begins with the activation of a region of chromatin. Active chromatin is decondensed. Decondensation appears to be accompanied by changes in binding coefficient,

phosphorylation and acetylation states of chromatin histones. HMG proteins HMG13 and HMG17 selectively bind activated chromatin. Topoisomerases remove superhelical tension on DNA. The activated region decondenses, allowing gene regulatory proteins and transcription factors to assemble on the DNA.

5        Patterns of chromatin structure can be stably inherited, producing heritable patterns of gene expression. In mammals, one of the two X chromosomes in each female cell is inactivated by condensation to heterochromatin during zygote development. The inactive state of this chromosome is inherited, so that adult females are mosaics of clusters of paternal-X and maternal-X clonal cell groups. The condensed X chromosome is reactivated in meiosis.

10       Chromatin is associated with disorders of protein expression such as thalassemia, a genetic anemia resulting from the removal of the locus control region (LCR) required for decondensation of the globin gene locus.

For a review of chromatin structure and function see Alberts, B. et al. (1994) Molecular Cell Biology, third edition, Garland Publishing, Inc., New York NY, pp. 351-354, 433-439.

15

#### Electron Transfer Associated Molecules

Electron carriers such as cytochromes accept electrons from NADH or FADH<sub>2</sub> and donate them to other electron carriers. Most electron-transferring proteins, except ubiquinone, are prosthetic groups such as flavins, heme, FeS clusters, and copper, bound to inner membrane proteins.

20       Adrenodoxin, for example, is an FeS protein that forms a complex with NADPH:adrenodoxin reductase and cytochrome p450. Cytochromes contain a heme prosthetic group, a porphyrin ring containing a tightly bound iron atom. Electron transfer reactions play a crucial role in cellular energy production.

25       Energy is produced by the oxidation of glucose and fatty acids. Glucose is initially converted to pyruvate in the cytoplasm. Fatty acids and pyruvate are transported to the mitochondria for complete oxidation to CO<sub>2</sub> coupled by enzymes to the transport of electrons from NADH and FADH<sub>2</sub> to oxygen and to the synthesis of ATP (oxidative phosphorylation) from ADP and P<sub>i</sub>.

Pyruvate is transported into the mitochondria and converted to acetyl-CoA for oxidation via the citric acid cycle, involving pyruvate dehydrogenase components, dihydrolipoyl transacetylase, and dihydrolipoyl dehydrogenase. Enzymes involved in the citric acid cycle include: citrate synthetase, aconitases, isocitrate dehydrogenase, alpha-ketoglutarate dehydrogenase complex including transsuccinylases, succinyl CoA synthetase, succinate dehydrogenase, fumarases, and malate dehydrogenase. Acetyl CoA is oxidized to CO<sub>2</sub> with concomitant formation of NADH, FADH<sub>2</sub>, and GTP. In oxidative phosphorylation, the transfer of electrons from NADH and FADH<sub>2</sub> to oxygen by dehydrogenases is coupled to the synthesis of ATP from ADP and P<sub>i</sub> by the F<sub>0</sub>F<sub>1</sub> ATPase complex in the mitochondrial inner membrane. Enzyme complexes responsible for electron transport

35

and ATP synthesis include the  $F_0F_1$  ATPase complex, ubiquinone(CoQ)-cytochrome c reductase, ubiquinone reductase, cytochrome b, cytochrome c<sub>1</sub>, FeS protein, and cytochrome c oxidase.

ATP synthesis requires membrane transport enzymes including the phosphate transporter and the ATP-ADP antiport protein. The ATP-binding cassette (ABC) superfamily has also been suggested  
5 as belonging to the mitochondrial transport group (Hogue, D.L. et al. (1999) J. Mol. Biol. 285:379-389). Brown fat uncoupling protein dissipates oxidative energy as heat, and may be involved the fever response to infection and trauma (Cannon, B. et al. (1998) Ann. NY Acad. Sci. 856:171-187).

Mitochondria are oval-shaped organelles comprising an outer membrane, a tightly folded inner membrane, an intermembrane space between the outer and inner membranes, and a matrix  
10 inside the inner membrane. The outer membrane contains many porin molecules that allow ions and charged molecules to enter the intermembrane space, while the inner membrane contains a variety of transport proteins that transfer only selected molecules. Mitochondria are the primary sites of energy production in cells.

Mitochondria contain a small amount of DNA. Human mitochondrial DNA encodes 13  
15 proteins, 22 tRNAs, and 2 rRNAs. Mitochondrial-DNA encoded proteins include NADH-Q reductase, a cytochrome reductase subunit, cytochrome oxidase subunits, and ATP synthase subunits.

Electron-transfer reactions also occur outside the mitochondria in locations such as the endoplasmic reticulum, which plays a crucial role in lipid and protein biosynthesis. Cytochrome b5 is a central electron donor for various reductive reactions occurring on the cytoplasmic surface of  
20 liver endoplasmic reticulum. Cytochrome b5 has been found in Golgi, plasma, endoplasmic reticulum (ER), and microbody membranes.

For a review of mitochondrial metabolism and regulation, see Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 745-797 and Stryer (1995) Biochemistry, W.H. Freeman and Co., San Francisco CA, pp 529-558, 988-989.

The majority of mitochondrial proteins are encoded by nuclear genes, are synthesized on cytosolic ribosomes, and are imported into the mitochondria. Nuclear-encoded proteins which are destined for the mitochondrial matrix typically contain positively-charged amino terminal signal sequences. Import of these preproteins from the cytoplasm requires a multisubunit protein complex in the outer membrane known as the translocase of outer mitochondrial membrane (TOM; previously  
30 designated MOM; Pfanner, N. et al. (1996) Trends Biochem. Sci. 21:51-52) and at least three inner membrane proteins which comprise the translocase of inner mitochondrial membrane (TIM; previously designated MIM; Pfanner, supra). An inside-negative membrane potential across the inner mitochondrial membrane is also required for preprotein import. Preproteins are recognized by surface receptor components of the TOM complex and are translocated through a proteinaceous pore  
35 formed by other TOM components. Proteins targeted to the matrix are then recognized by the import machinery of the TIM complex. The import systems of the outer and inner membranes can function

independently (Segui-Real, B. et al. (1993) EMBO J. 12:2211-2218).

Once precursor proteins are in the mitochondria, the leader peptide is cleaved by a signal peptidase to generate the mature protein. Most leader peptides are removed in a one step process by a protease termed mitochondrial processing peptidase (MPP) (Paces, V. et al. (1993) Proc. Natl. Acad. Sci. USA 90:5355-5358). In some cases a two-step process occurs in which MPP generates an intermediate precursor form which is cleaved by a second enzyme, mitochondrial intermediate peptidase, to generate the mature protein.

Mitochondrial dysfunction leads to impaired calcium buffering, generation of free radicals that may participate in deleterious intracellular and extracellular processes, changes in mitochondrial permeability and oxidative damage which is observed in several neurodegenerative diseases. Neurodegenerative diseases linked to mitochondrial dysfunction include some forms of Alzheimer's disease, Friedreich's ataxia, familial amyotrophic lateral sclerosis, and Huntington's disease (Beal, M.F. (1998) Biochim. Biophys. Acta 1366:211-213). The myocardium is heavily dependent on oxidative metabolism, so mitochondrial dysfunction often leads to heart disease (DiMauro, S. and M. Hirano (1998) Curr. Opin. Cardiol 13:190-197). Mitochondria are implicated in disorders of cell proliferation, since they play an important role in a cell's decision to proliferate or self-destruct through apoptosis. The oncoprotein Bcl-2, for example, promotes cell proliferation by stabilizing mitochondrial membranes so that apoptosis signals are not released (Susin, S.A. (1998) Biochim. Biophys. Acta 1366:151-165).

#### **Transcription Factor Molecules**

Multicellular organisms are comprised of diverse cell types that differ dramatically both in structure and function. The identity of a cell is determined by its characteristic pattern of gene expression, and different cell types express overlapping but distinctive sets of genes throughout development. Spatial and temporal regulation of gene expression is critical for the control of cell proliferation, cell differentiation, apoptosis, and other processes that contribute to organismal development. Furthermore, gene expression is regulated in response to extracellular signals that mediate cell-cell communication and coordinate the activities of different cell types. Appropriate gene regulation also ensures that cells function efficiently by expressing only those genes whose functions are required at a given time.

Transcriptional regulatory proteins are essential for the control of gene expression. Some of these proteins function as transcription factors that initiate, activate, repress, or terminate gene transcription. Transcription factors generally bind to the promoter, enhancer, and upstream regulatory regions of a gene in a sequence-specific manner, although some factors bind regulatory elements within or downstream of a gene's coding region. Transcription factors may bind to a specific region of DNA singly or as a complex with other accessory factors. (Reviewed in Lewin, B.

(1990) Genes IV, Oxford University Press, New York NY, and Cell Press, Cambridge MA, pp. 554-570.)

The double helix structure and repeated sequences of DNA create topological and chemical features which can be recognized by transcription factors. These features are hydrogen bond donor and acceptor groups, hydrophobic patches, major and minor grooves, and regular, repeated stretches of sequence which induce distinct bends in the helix. Typically, transcription factors recognize specific DNA sequence motifs of about 20 nucleotides in length. Multiple, adjacent transcription factor-binding motifs may be required for gene regulation.

Many transcription factors incorporate DNA-binding structural motifs which comprise either  $\alpha$  helices or  $\beta$  sheets that bind to the major groove of DNA. Four well-characterized structural motifs are helix-turn-helix, zinc finger, leucine zipper, and helix-loop-helix. Proteins containing these motifs may act alone as monomers, or they may form homo- or heterodimers that interact with DNA.

The helix-turn-helix motif consists of two  $\alpha$  helices connected at a fixed angle by a short chain of amino acids. One of the helices binds to the major groove. Helix-turn-helix motifs are exemplified by the homeobox motif which is present in homeodomain proteins. These proteins are critical for specifying the anterior-posterior body axis during development and are conserved throughout the animal kingdom. The Antennapedia and Ultrabithorax proteins of Drosophila melanogaster are prototypical homeodomain proteins (Pabo, C.O. and R.T. Sauer (1992) Annu. Rev. Biochem. 61:1053-1095).

The zinc finger motif, which binds zinc ions, generally contains tandem repeats of about 30 amino acids consisting of periodically spaced cysteine and histidine residues. Examples of this sequence pattern, designated C2H2 and C3HC4 ("RING" finger), have been described (Lewin, supra). Zinc finger proteins each contain an  $\alpha$  helix and an antiparallel  $\beta$  sheet whose proximity and conformation are maintained by the zinc ion. Contact with DNA is made by the arginine preceding the  $\alpha$  helix and by the second, third, and sixth residues of the  $\alpha$  helix. Variants of the zinc finger motif include poorly defined cysteine-rich motifs which bind zinc or other metal ions. These motifs may not contain histidine residues and are generally nonrepetitive.

The leucine zipper motif comprises a stretch of amino acids rich in leucine which can form an amphipathic  $\alpha$  helix. This structure provides the basis for dimerization of two leucine zipper proteins. The region adjacent to the leucine zipper is usually basic, and upon protein dimerization, is optimally positioned for binding to the major groove. Proteins containing such motifs are generally referred to as bZIP transcription factors.

The helix-loop-helix motif (HLH) consists of a short  $\alpha$  helix connected by a loop to a longer  $\alpha$  helix. The loop is flexible and allows the two helices to fold back against each other and to bind to DNA. The transcription factor Myc contains a prototypical HLH motif.

Most transcription factors contain characteristic DNA binding motifs, and variations on the above motifs and new motifs have been and are currently being characterized (Faisst, S. and S. Meyer (1992) *Nucleic Acids Res.* 20:3-26).

Many neoplastic disorders in humans can be attributed to inappropriate gene expression.

5 Malignant cell growth may result from either excessive expression of tumor promoting genes or insufficient expression of tumor suppressor genes (Cleary, M.L. (1992) *Cancer Surv.* 15:89-104). Chromosomal translocations may also produce chimeric loci which fuse the coding sequence of one gene with the regulatory regions of a second unrelated gene. Such an arrangement likely results in inappropriate gene transcription, potentially contributing to malignancy.

10 In addition, the immune system responds to infection or trauma by activating a cascade of events that coordinate the progressive selection, amplification, and mobilization of cellular defense mechanisms. A complex and balanced program of gene activation and repression is involved in this process. However, hyperactivity of the immune system as a result of improper or insufficient regulation of gene expression may result in considerable tissue or organ damage. This damage is  
15 well documented in immunological responses associated with arthritis, allergens, heart attack, stroke, and infections (Isselbacher, K.J. et al. (1996) Harrison's Principles of Internal Medicine, 13/e, McGraw Hill, Inc. and Teton Data Systems Software).

Furthermore, the generation of multicellular organisms is based upon the induction and coordination of cell differentiation at the appropriate stages of development. Central to this process  
20 is differential gene expression, which confers the distinct identities of cells and tissues throughout the body. Failure to regulate gene expression during development can result in developmental disorders. Human developmental disorders caused by mutations in zinc finger-type transcriptional regulators include: urogenital developmental abnormalities associated with WT1; Greig cephalopolysyndactyly, Pallister-Hall syndrome, and postaxial polydactyly type A (GLI3); and  
25 Townes-Brocks syndrome, characterized by anal, renal, limb, and ear abnormalities (SALL1) (Engelkamp, D. and V. van Heyningen (1996) *Curr. Opin. Genet. Dev.* 6:334-342; Kohlhase, J. et al. (1999) *Am. J. Hum. Genet.* 64:435-445).

### Cell Membrane Molecules

30 Eukaryotic cells are surrounded by plasma membranes which enclose the cell and maintain an environment inside the cell that is distinct from its surroundings. In addition, eukaryotic organisms are distinct from prokaryotes in possessing many intracellular organelle and vesicle structures. Many of the metabolic reactions which distinguish eukaryotic biochemistry from prokaryotic biochemistry take place within these structures. The plasma membrane and the  
35 membranes surrounding organelles and vesicles are composed of phosphoglycerides, fatty acids, cholesterol, phospholipids, glycolipids, proteoglycans, and proteins. These components confer

identity and functionality to the membranes with which they associate.

### Integral Membrane Proteins

The majority of known integral membrane proteins are transmembrane proteins (TM) which are characterized by an extracellular, a transmembrane, and an intracellular domain. TM domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an  $\alpha$ -helical conformation. TM proteins are classified as bitopic (Types I and II) and polytopic (Types III and IV) (Singer, S.J. (1990) *Annu. Rev. Cell Biol.* 6:247-296). Bitopic proteins span the membrane once while polytopic proteins contain multiple membrane-spanning segments. TM proteins function as cell-surface receptors, receptor-interacting proteins, transporters of ions or metabolites, ion channels, cell anchoring proteins, and cell type-specific surface antigens.

Many membrane proteins (MPs) contain amino acid sequence motifs that target these proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in targeted cancer treatment of tumor vasculature (Arap, W. et al. (1998) *Science* 279:377-380). Furthermore, MPs may also contain amino acid sequence motifs, such as the carbohydrate recognition domain (CRD), that mediate interactions with extracellular or intracellular molecules.

### G-Protein Coupled Receptors

G-protein coupled receptors (GPCR) are a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines, lipid mediators of inflammation, peptide hormones, and sensory signal mediators. The structure of these highly-conserved receptors consists of seven hydrophobic transmembrane regions, an extracellular N-terminus, and a cytoplasmic C-terminus. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. Cysteine disulfide bridges connect the second and third extracellular loops. The most conserved regions of GPCRs are the transmembrane regions and the first two cytoplasmic loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with G proteins. A GPCR consensus pattern is characteristic of most proteins belonging to this superfamily (ExPASy PROSITE document PS00237; and Watson, S. and S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 2-6). Mutations and changes in transcriptional activation of GPCR-encoding genes have been associated with neurological disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, drug addiction, and feeding disorders.

### Scavenger Receptors

Macrophage scavenger receptors with broad ligand specificity may participate in the binding of low density lipoproteins (LDL) and foreign antigens. Scavenger receptors types I and II are trimeric membrane proteins with each subunit containing a small N-terminal intracellular domain, a

transmembrane domain, a large extracellular domain, and a C-terminal cysteine-rich domain. The extracellular domain contains a short spacer region, an  $\alpha$ -helical coiled-coil region, and a triple helical collagen-like region. These receptors have been shown to bind a spectrum of ligands, including chemically modified lipoproteins and albumin, polyribonucleotides, polysaccharides, phospholipids, and asbestos (Matsumoto, A. et al. (1990) Proc. Natl. Acad. Sci. USA 87:9133-9137; and Elomaa, O. et al. (1995) Cell 80:603-609). The scavenger receptors are thought to play a key role in atherogenesis by mediating uptake of modified LDL in arterial walls, and in host defense by binding bacterial endotoxins, bacteria, and protozoa.

#### Tetraspan Family Proteins

10 The transmembrane 4 superfamily (TM4SF) or tetraspan family is a multigene family encoding type III integral membrane proteins (Wright, M.D. and M.G. Tomlinson (1994) Immunol. Today 15:588-594). The TM4SF is comprised of membrane proteins which traverse the cell membrane four times. Members of the TM4SF include platelet and endothelial cell membrane proteins, melanoma-associated antigens, leukocyte surface glycoproteins, colonic carcinoma  
15 antigens, tumor-associated antigens, and surface proteins of the schistosome parasites (Jankowski, S.A. (1994) Oncogene 9:1205-1211). Members of the TM4SF share about 25-30% amino acid sequence identity with one another.

A number of TM4SF members have been implicated in signal transduction, control of cell adhesion, regulation of cell growth and proliferation, including development and oncogenesis, and  
20 cell motility, including tumor cell metastasis. Expression of TM4SF proteins is associated with a variety of tumors and the level of expression may be altered when cells are growing or activated.

#### Tumor Antigens

Tumor antigens are cell surface molecules that are differentially expressed in tumor cells relative to normal cells. Tumor antigens distinguish tumor cells immunologically from normal cells  
25 and provide diagnostic and therapeutic targets for human cancers (Takagi, S. et al. (1995) Int. J. Cancer 61:706-715; Liu, E. et al. (1992) Oncogene 7:1027-1032).

#### Leukocyte Antigens

Other types of cell surface antigens include those identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-  
30 based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD"  
35 designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both



transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI).

(Reviewed in Barclay, A.N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego CA, pp. 17-20.)

## 5 Ion Channels

Ion channels are found in the plasma membranes of virtually every cell in the body. For example, chloride channels mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of ions across epithelial membranes. Chloride channels also regulate the pH of organelles such as the Golgi apparatus and endosomes (see, e.g., Greger, R. (1988) 10 *Annu. Rev. Physiol.* 50:111-122). Electrophysiological and pharmacological properties of chloride channels, including ion conductance, current-voltage relationships, and sensitivity to modulators, suggest that different chloride channels exist in muscles, neurons, fibroblasts, epithelial cells, and lymphocytes.

Many ion channels have sites for phosphorylation by one or more protein kinases including 15 protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Inappropriate phosphorylation of proteins in cells has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, skeletal muscle, and other organ systems.

## 20 Proton Pumps

Proton ATPases comprise a large class of membrane proteins that use the energy of ATP hydrolysis to generate an electrochemical proton gradient across a membrane. The resultant gradient may be used to transport other ions across the membrane ( $\text{Na}^+$ ,  $\text{K}^+$ , or  $\text{Cl}^-$ ) or to maintain organelle pH. Proton ATPases are further subdivided into the mitochondrial F-ATPases, the plasma membrane 25 ATPases, and the vacuolar ATPases. The vacuolar ATPases establish and maintain an acidic pH within various organelles involved in the processes of endocytosis and exocytosis (Mellman, I. et al. (1986) *Annu. Rev. Biochem.* 55:663-700).

Proton-coupled, 12 membrane-spanning domain transporters such as PEPT 1 and PEPT 2 are responsible for gastrointestinal absorption and for renal reabsorption of peptides using an 30 electrochemical  $\text{H}^+$  gradient as the driving force. Another type of peptide transporter, the TAP transporter, is a heterodimer consisting of TAP 1 and TAP 2 and is associated with antigen processing. Peptide antigens are transported across the membrane of the endoplasmic reticulum by TAP so they can be expressed on the cell surface in association with MHC molecules. Each TAP protein consists of multiple hydrophobic membrane spanning segments and a highly conserved 35 ATP-binding cassette (Boll, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:284-289). Pathogenic microorganisms, such as herpes simplex virus, may encode inhibitors of TAP-mediated peptide

transport in order to evade immune surveillance (Marusina, K. and J.J Manaco (1996) Curr. Opin. Hematol. 3:19-26).

#### ABC Transporters

The ATP-binding cassette (ABC) transporters, also called the "traffic ATPases", comprise a  
5 superfamily of membrane proteins that mediate transport and channel functions in prokaryotes and  
eukaryotes (Higgins, C.F. (1992) Annu. Rev. Cell Biol. 8:67-113). ABC proteins share a similar  
overall structure and significant sequence homology. All ABC proteins contain a conserved domain  
of approximately two hundred amino acid residues which includes one or more nucleotide binding  
domains. Mutations in ABC transporter genes are associated with various disorders, such as  
10 hyperbilirubinemia II/Dubin-Johnson syndrome, recessive Stargardt's disease, X-linked  
adrenoleukodystrophy, multidrug resistance, celiac disease, and cystic fibrosis.

#### Peripheral and Anchored Membrane Proteins

Some membrane proteins are not membrane-spanning but are attached to the plasma  
membrane via membrane anchors or interactions with integral membrane proteins. Membrane  
15 anchors are covalently joined to a protein post-translationally and include such moieties as prenyl,  
myristyl, and glycosylphosphatidyl inositol groups. Membrane localization of peripheral and  
anchored proteins is important for their function in processes such as receptor-mediated signal  
transduction. For example, prenylation of Ras is required for its localization to the plasma membrane  
and for its normal and oncogenic functions in signal transduction.

#### Vesicle Coat Proteins

Intercellular communication is essential for the development and survival of multicellular  
organisms. Cells communicate with one another through the secretion and uptake of protein  
signaling molecules. The uptake of proteins into the cell is achieved by the endocytic pathway, in  
which the interaction of extracellular signaling molecules with plasma membrane receptors results in  
25 the formation of plasma membrane-derived vesicles that enclose and transport the molecules into the  
cytosol. These transport vesicles fuse with and mature into endosomal and lysosomal (digestive)  
compartments. The secretion of proteins from the cell is achieved by exocytosis, in which molecules  
inside of the cell proceed through the secretory pathway. In this pathway, molecules transit from the  
ER to the Golgi apparatus and finally to the plasma membrane, where they are secreted from the cell.

30 Several steps in the transit of material along the secretory and endocytic pathways require the  
formation of transport vesicles. Specifically, vesicles form at the transitional endoplasmic reticulum  
(tER), the rim of Golgi cisternae, the face of the Trans-Golgi Network (TGN), the plasma membrane  
(PM), and tubular extensions of the endosomes. Vesicle formation occurs when a region of  
membrane buds off from the donor organelle. The membrane-bound vesicle contains proteins to be  
35 transported and is surrounded by a proteinaceous coat, the components of which are recruited from  
the cytosol. Two different classes of coat protein have been identified. Clathrin coats form on

vesicles derived from the TGN and PM, whereas coatomer (COP) coats form on vesicles derived from the ER and Golgi. COP coats can be further classified as COPI, involved in retrograde traffic through the Golgi and from the Golgi to the ER, and COPII, involved in anterograde traffic from the ER to the Golgi (Mellman, *supra*).

5 In clathrin-based vesicle formation, adapter proteins bring vesicle cargo and coat proteins together at the surface of the budding membrane. Adapter protein-1 and -2 select cargo from the TGN and plasma membrane, respectively, based on molecular information encoded on the cytoplasmic tail of integral membrane cargo proteins. Adapter proteins also recruit clathrin to the bud site. Clathrin is a protein complex consisting of three large and three small polypeptide chains  
10 arranged in a three-legged structure called a triskelion. Multiple triskelions and other coat proteins appear to self-assemble on the membrane to form a coated pit. This assembly process may serve to deform the membrane into a budding vesicle. GTP-bound ADP-ribosylation factor (Arf) is also incorporated into the coated assembly. Another small G-protein, dynamin, forms a ring complex around the neck of the forming vesicle and may provide the mechanochemical force to seal the bud,  
15 thereby releasing the vesicle. The coated vesicle complex is then transported through the cytosol. During the transport process, Arf-bound GTP is hydrolyzed to GDP, and the coat dissociates from the transport vesicle (West, M.A. et al. (1997) J. Cell Biol. 138:1239-1254).

Vesicles which bud from the ER and the Golgi are covered with a protein coat similar to the clathrin coat of endocytic and TGN vesicles. The coat protein (COP) is assembled from cytosolic  
20 precursor molecules at specific budding regions on the organelle. The COP coat consists of two major components, a G-protein (Arf or Sar) and coat protomer (coatomer). Coatomer is an equimolar complex of seven proteins, termed alpha-, beta-, beta'-, gamma-, delta-, epsilon- and zeta-COP. The coatomer complex binds to dilysine motifs contained on the cytoplasmic tails of integral membrane proteins. These include the KKXX retrieval motif of membrane proteins of the ER and  
25 dibasic/diphenylamine motifs of members of the p24 family. The p24 family of type I membrane proteins represent the major membrane proteins of COPI vesicles (Harter, C. and F.T. Wieland (1998) Proc. Natl. Acad. Sci. USA 95:11649-11654).

### Organelle Associated Molecules

30 Eukaryotic cells are organized into various cellular organelles which has the effect of separating specific molecules and their functions from one another and from the cytosol. Within the cell, various membrane structures surround and define these organelles while allowing them to interact with one another and the cell environment through both active and passive transport processes. Important cell organelles include the nucleus, the Golgi apparatus, the endoplasmic  
35 reticulum, mitochondria, peroxisomes, lysosomes, endosomes, and secretory vesicles.

### Nucleus

The cell nucleus contains all of the genetic information of the cell in the form of DNA, and the components and machinery necessary for replication of DNA and for transcription of DNA into RNA. (See Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing Inc., New York NY, pp. 335-399.) DNA is organized into compact structures in the nucleus by interactions  
5 with various DNA-binding proteins such as histones and non-histone chromosomal proteins. DNA-specific nucleases, DNAses, partially degrade these compacted structures prior to DNA replication or transcription. DNA replication takes place with the aid of DNA helicases which unwind the double-stranded DNA helix, and DNA polymerases that duplicate the separated DNA strands.

10 Transcriptional regulatory proteins are essential for the control of gene expression. Some of these proteins function as transcription factors that initiate, activate, repress, or terminate gene transcription. Transcription factors generally bind to the promoter, enhancer, and upstream regulatory regions of a gene in a sequence-specific manner, although some factors bind regulatory elements within or downstream of a gene's coding region. Transcription factors may bind to a  
15 specific region of DNA singly or as a complex with other accessory factors. (Reviewed in Lewin, B. (1990) Genes IV, Oxford University Press, New York NY, and Cell Press, Cambridge MA, pp. 554-570.) Many transcription factors incorporate DNA-binding structural motifs which comprise either  $\alpha$  helices or  $\beta$  sheets that bind to the major groove of DNA. Four well-characterized structural motifs are helix-turn-helix, zinc finger, leucine zipper, and helix-loop-helix. Proteins containing  
20 these motifs may act alone as monomers, or they may form homo- or heterodimers that interact with DNA.

Many neoplastic disorders in humans can be attributed to inappropriate gene expression. Malignant cell growth may result from either excessive expression of tumor promoting genes or insufficient expression of tumor suppressor genes (Cleary, M.L. (1992) *Cancer Surv.* 15:89-104).  
25 Chromosomal translocations may also produce chimeric loci which fuse the coding sequence of one gene with the regulatory regions of a second unrelated gene. Such an arrangement likely results in inappropriate gene transcription, potentially contributing to malignancy.

In addition, the immune system responds to infection or trauma by activating a cascade of events that coordinate the progressive selection, amplification, and mobilization of cellular defense  
30 mechanisms. A complex and balanced program of gene activation and repression is involved in this process. However, hyperactivity of the immune system as a result of improper or insufficient regulation of gene expression may result in considerable tissue or organ damage. This damage is well documented in immunological responses associated with arthritis, allergens, heart attack, stroke, and infections (Isselbacher, K.J. et al. (1996) Harrison's Principles of Internal Medicine, 13/e,  
35 McGraw Hill, Inc. and Teton Data Systems Software).

Transcription of DNA into RNA also takes place in the nucleus catalyzed by RNA

polymerases. Three types of RNA polymerase exist. RNA polymerase I makes large ribosomal RNAs, while RNA polymerase III makes a variety of small, stable RNAs including 5S ribosomal RNA and the transfer RNAs (tRNA). RNA polymerase II transcribes genes that will be translated into proteins. The primary transcript of RNA polymerase II is called heterogeneous nuclear RNA (hnRNA), and must be further processed by splicing to remove non-coding sequences called introns. RNA splicing is mediated by small nuclear ribonucleoprotein complexes, or snRNPs, producing mature messenger RNA (mRNA) which is then transported out of the nucleus for translation into proteins.

### Nucleolus

The nucleolus is a highly organized subcompartment in the nucleus that contains high concentrations of RNA and proteins and functions mainly in ribosomal RNA synthesis and assembly (Alberts, et al. *supra*, pp. 379-382). Ribosomal RNA (rRNA) is a structural RNA that is complexed with proteins to form ribonucleoprotein structures called ribosomes. Ribosomes provide the platform on which protein synthesis takes place.

Ribosomes are assembled in the nucleolus initially from a large, 45S rRNA combined with a variety of proteins imported from the cytoplasm, as well as smaller, 5S rRNAs. Later processing of the immature ribosome results in formation of smaller ribosomal subunits which are transported from the nucleolus to the cytoplasm where they are assembled into functional ribosomes.

### Endoplasmic Reticulum

In eukaryotes, proteins are synthesized within the endoplasmic reticulum (ER), delivered from the ER to the Golgi apparatus for post-translational processing and sorting, and transported from the Golgi to specific intracellular and extracellular destinations. Synthesis of integral membrane proteins, secreted proteins, and proteins destined for the lumen of a particular organelle occurs on the rough endoplasmic reticulum (ER). The rough ER is so named because of the rough appearance in electron micrographs imparted by the attached ribosomes on which protein synthesis proceeds. Synthesis of proteins destined for the ER actually begins in the cytosol with the synthesis of a specific signal peptide which directs the growing polypeptide and its attached ribosome to the ER membrane where the signal peptide is removed and protein synthesis is completed. Soluble proteins destined for the ER lumen, for secretion, or for transport to the lumen of other organelles pass completely into the ER lumen. Transmembrane proteins destined for the ER or for other cell membranes are translocated across the ER membrane but remain anchored in the lipid bilayer of the membrane by one or more membrane-spanning  $\alpha$ -helical regions.

Translocated polypeptide chains destined for other organelles or for secretion also fold and assemble in the ER lumen with the aid of certain "resident" ER proteins. Protein folding in the ER is aided by two principal types of protein isomerases, protein disulfide isomerase (PDI), and peptidyl-prolyl isomerase (PPI). PDI catalyzes the oxidation of free sulfhydryl groups in cysteine residues to

form intramolecular disulfide bonds in proteins. PPI, an enzyme that catalyzes the isomerization of certain proline imide bonds in oligopeptides and proteins, is considered to govern one of the rate limiting steps in the folding of many proteins to their final functional conformation. The cyclophilins represent a major class of PPI that was originally identified as the major receptor for the

5 immunosuppressive drug cyclosporin A (Handschumacher, R.E. et al. (1984) Science 226:544-547). Molecular "chaperones" such as BiP (binding protein) in the ER recognize incorrectly folded proteins as well as proteins not yet folded into their final form and bind to them, both to prevent improper aggregation between them, and to promote proper folding.

The "N-linked" glycosylation of most soluble secreted and membrane-bound proteins by  
10 oligosacchrides linked to asparagine residues in proteins is also performed in the ER. This reaction is catalyzed by a membrane-bound enzyme, oligosaccharyl transferase.

#### Golgi Apparatus

The Golgi apparatus is a complex structure that lies adjacent to the ER in eukaryotic cells and serves primarily as a sorting and dispatching station for products of the ER (Alberts, et al. supra, pp.  
15 600-610). Additional posttranslational processing, principally additional glycosylation, also occurs in the Golgi. Indeed, the Golgi is a major site of carbohydrate synthesis, including most of the glycosaminoglycans of the extracellular matrix. N-linked oligosaccharides, added to proteins in the ER, are also further modified in the Golgi by the addition of more sugar residues to form complex N-linked oligosaccharides. "O-linked" glycosylation of proteins also occurs in the Golgi by the addition  
20 of N-acetylgalactosamine to the hydroxyl group of a serine or threonine residue followed by the sequential addition of other sugar residues to the first. This process is catalyzed by a series of glycosyltransferases each specific for a particular donor sugar nucleotide and acceptor molecule (Lodish, H. et al. (1995) Molecular Cell Biology, W.H. Freeman and Co., New York NY, pp.700-708). In many cases, both N- and O-linked oligosaccharides appear to be required for the secretion of  
25 proteins or the movement of plasma membrane glycoproteins to the cell surface.

The terminal compartment of the Golgi is the Trans-Golgi Network (TGN), where both membrane and lumenal proteins are sorted for their final destination. Transport (or secretory) vesicles destined for intracellular compartments, such as lysosomes, bud off of the TGN. Other transport vesicles bud off containing proteins destined for the plasma membrane, such as receptors,  
30 adhesion molecules, and ion channels, and secretory proteins, such as hormones, neurotransmitters, and digestive enzymes.

#### Vacuoles

The vacuole system is a collection of membrane bound compartments in eukaryotic cells that functions in the processes of endocytosis and exocytosis. They include phagosomes, lysosomes,  
35 endosomes, and secretory vesicles. Endocytosis is the process in cells of internalizing nutrients, solutes or small particles (pinocytosis) or large particles such as internalized receptors, viruses,

bacteria, or bacterial toxins (phagocytosis). Exocytosis is the process of transporting molecules to the cell surface. It facilitates placement or localization of membrane-bound receptors or other membrane proteins and secretion of hormones, neurotransmitters, digestive enzymes, wastes, etc.

A common property of all of these vacuoles is an acidic pH environment ranging from approximately pH 4.5-5.0. This acidity is maintained by the presence of a proton ATPase that uses the energy of ATP hydrolysis to generate an electrochemical proton gradient across a membrane (Mellman, I. et al. (1986) *Annu. Rev. Biochem.* 55:663-700). Eukaryotic vacuolar proton ATPase (vp-ATPase) is a multimeric enzyme composed of 3-10 different subunits. One of these subunits is a highly hydrophobic polypeptide of approximately 16 kDa that is similar to the proteolipid component of vp-ATPases from eubacteria, fungi, and plant vacuoles (Mandel, M. et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:5521-5524). The 16 kDa proteolipid component is the major subunit of the membrane portion of vp-ATPase and functions in the transport of protons across the membrane.

#### Lysosomes

Lysosomes are membranous vesicles containing various hydrolytic enzymes used for the controlled intracellular digestion of macromolecules. Lysosomes contain some 40 types of enzymes including proteases, nucleases, glycosidases, lipases, phospholipases, phosphatases, and sulfatases, all of which are acid hydrolases that function at a pH of about 5. Lysosomes are surrounded by a unique membrane containing transport proteins that allow the final products of macromolecule degradation, such as sugars, amino acids, and nucleotides, to be transported to the cytosol where they may be either excreted or reutilized by the cell. A vp-ATPase, such as that described above, maintains the acidic environment necessary for hydrolytic activity (Alberts, *supra*, pp. 610-611).

#### Endosomes

Endosomes are another type of acidic vacuole that is used to transport substances from the cell surface to the interior of the cell in the process of endocytosis. Like lysosomes, endosomes have an acidic environment provided by a vp-ATPase (Alberts et al. *supra*, pp. 610-618). Two types of endosomes are apparent based on tracer uptake studies that distinguish their time of formation in the cell and their cellular location. Early endosomes are found near the plasma membrane and appear to function primarily in the recycling of internalized receptors back to the cell surface. Late endosomes appear later in the endocytic process close to the Golgi apparatus and the nucleus, and appear to be associated with delivery of endocytosed material to lysosomes or to the TGN where they may be recycled. Specific proteins are associated with particular transport vesicles and their target compartments that may provide selectivity in targeting vesicles to their proper compartments. A cytosolic prenylated GTP-binding protein, Rab, is one such protein. Rabs 4, 5, and 11 are associated with the early endosome, whereas Rabs 7 and 9 associate with the late endosome.

#### Mitochondria

Mitochondria are oval-shaped organelles comprising an outer membrane, a tightly folded

inner membrane, an intermembrane space between the outer and inner membranes, and a matrix inside the inner membrane. The outer membrane contains many porin molecules that allow ions and charged molecules to enter the intermembrane space, while the inner membrane contains a variety of transport proteins that transfer only selected molecules. Mitochondria are the primary sites of energy production in cells.

Energy is produced by the oxidation of glucose and fatty acids. Glucose is initially converted to pyruvate in the cytoplasm. Fatty acids and pyruvate are transported to the mitochondria for complete oxidation to  $\text{CO}_2$  coupled by enzymes to the transport of electrons from NADH and  $\text{FADH}_2$  to oxygen and to the synthesis of ATP (oxidative phosphorylation) from ADP and  $\text{P}_i$ .

Pyruvate is transported into the mitochondria and converted to acetyl-CoA for oxidation via the citric acid cycle, involving pyruvate dehydrogenase components, dihydrolipoyl transacetylase, and dihydrolipoyl dehydrogenase. Enzymes involved in the citric acid cycle include: citrate synthetase, aconitases, isocitrate dehydrogenase, alpha-ketoglutarate dehydrogenase complex including transsuccinylases, succinyl CoA synthetase, succinate dehydrogenase, fumarases, and malate dehydrogenase. Acetyl CoA is oxidized to  $\text{CO}_2$  with concomitant formation of NADH,  $\text{FADH}_2$ , and GTP. In oxidative phosphorylation, the transfer of electrons from NADH and  $\text{FADH}_2$  to oxygen by dehydrogenases is coupled to the synthesis of ATP from ADP and  $\text{P}_i$  by the  $\text{F}_0\text{F}_1$  ATPase complex in the mitochondrial inner membrane. Enzyme complexes responsible for electron transport and ATP synthesis include the  $\text{F}_0\text{F}_1$  ATPase complex, ubiquinone(CoQ)-cytochrome c reductase, ubiquinone reductase, cytochrome b, cytochrome c, FeS protein, and cytochrome c oxidase.

#### Peroxisomes

Peroxisomes, like mitochondria, are a major site of oxygen utilization. They contain one or more enzymes, such as catalase and urate oxidase, that use molecular oxygen to remove hydrogen atoms from specific organic substrates in an oxidative reaction that produces hydrogen peroxide (Alberts, *supra*, pp. 574-577). Catalase oxidizes a variety of substrates including phenols, formic acid, formaldehyde, and alcohol and is important in peroxisomes of liver and kidney cells for detoxifying various toxic molecules that enter the bloodstream. Another major function of oxidative reactions in peroxisomes is the breakdown of fatty acids in a process called  $\beta$  oxidation.  $\beta$  oxidation results in shortening of the alkyl chain of fatty acids by blocks of two carbon atoms that are converted to acetyl CoA and exported to the cytosol for reuse in biosynthetic reactions.

Also like mitochondria, peroxisomes import their proteins from the cytosol using a specific signal sequence located near the C-terminus of the protein. The importance of this import process is evident in the inherited human disease Zellweger syndrome, in which a defect in importing proteins into peroxisomes leads to a peroxisomal deficiency resulting in severe abnormalities in the brain, liver, and kidneys, and death soon after birth. One form of this disease has been shown to be due to a mutation in the gene encoding a peroxisomal integral membrane protein called peroxisome assembly



factor-1.

The discovery of new human molecules satisfies a need in the art by providing new compositions which are useful in the diagnosis, study, prevention, and treatment of diseases associated with, as well as effects of exogenous compounds on, the expression of human molecules.

5

### SUMMARY OF THE INVENTION

The present invention relates to nucleic acid sequences comprising human diagnostic and therapeutic polynucleotides (dithp) as presented in the Sequence Listing. The dithp uniquely identify genes encoding human structural, functional, and regulatory molecules.

10 The invention provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). In one alternative, the polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275. In another alternative, the polynucleotide comprises at least 30 contiguous nucleotides of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). In another alternative, the polynucleotide comprises at least 60 contiguous nucleotides of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). The invention further provides a composition for the detection of expression of human diagnostic and therapeutic polynucleotides comprising at least one isolated polynucleotide comprising a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a

15

20

25

30

35

polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d); and a detectable label.

The invention also provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a polynucleotide sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence of a polynucleotide selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention also provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a polynucleotide sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof. In one alternative, the invention provides a composition comprising a target polynucleotide of the method, wherein said probe comprises at least 30 contiguous nucleotides. In one alternative, the invention provides a composition comprising a target polynucleotide of the method, wherein said probe comprises at least 60 contiguous nucleotides.

The invention further provides a recombinant polynucleotide comprising a promoter sequence operably linked to an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the

invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a human diagnostic and therapeutic polypeptide, the method comprising a) culturing a cell under conditions suitable for expression of the human diagnostic and therapeutic polypeptide, wherein said cell is transformed with a recombinant polynucleotide, said recombinant polynucleotide comprising an isolated polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; iii) a polynucleotide complementary to the polynucleotide of i); iv) a polynucleotide complementary to the polynucleotide of ii); and v) an RNA equivalent of i) through iv), and b) recovering the human diagnostic and therapeutic polypeptide so expressed. The invention additionally provides a method wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

The invention also provides an isolated human diagnostic and therapeutic polypeptide (DITHP) encoded by at least one polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275. The invention further provides a method of screening for a test compound that specifically binds to the polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. The method comprises a) combining the polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553 with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553 to the test compound, thereby identifying a compound that specifically binds to the polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

The invention further provides a microarray wherein at least one element of the microarray is an isolated polynucleotide comprising at least 30 contiguous nucleotides of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). The invention also provides a method for generating a transcript image of a sample which contains polynucleotides. The method comprises a) labeling the polynucleotides of the sample, b) contacting the elements of the microarray with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and c) quantifying the expression of the polynucleotides in the sample.

Additionally, the invention provides a method for screening a compound for effectiveness in

altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence  
5 selected from the group consisting of SEQ ID NO:1-275; c) a polynucleotide complementary to the polynucleotide of a); d) a polynucleotide complementary to the polynucleotide of b); and e) an RNA equivalent of a) through d). The method comprises a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the  
10 compound and in the absence of the compound.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide  
15 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; iii) a polynucleotide complementary to the polynucleotide of i); iv) a polynucleotide complementary to the polynucleotide of ii); and v) an RNA equivalent of i) through iv). Hybridization occurs under  
20 conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a  
25 polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275; iii) a polynucleotide complementary to the polynucleotide of i); iv) a polynucleotide complementary to the polynucleotide of ii); and v) an RNA equivalent of i) through iv), and alternatively, the target polynucleotide comprises a polynucleotide sequence of a fragment of a polynucleotide selected from the group consisting of i-v above; c) quantifying the amount of hybridization complex; and d)  
30 comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

The invention further provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID  
35 NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a

biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. In one alternative, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. In one alternative, the polynucleotide encodes a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. In another alternative, the polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

The invention further provides a composition comprising a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional DITHP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional DITHP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional DITHP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c)

comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

5

### DESCRIPTION OF THE TABLES

Table 1 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with the sequence identification numbers (SEQ ID NO:s) and open reading frame identification  
10 numbers (ORF IDs) corresponding to polypeptides encoded by the template ID.

Table 2 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with their GenBank hits (GI Numbers), probability scores, and functional annotations corresponding to the GenBank hits.

15 Table 3 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with polynucleotide segments of each template sequence as defined by the indicated "start" and "stop" nucleotide positions. The reading frames of the polynucleotide segments and the Pfam hits, Pfam descriptions, and E-values corresponding to the polypeptide domains encoded by the  
20 polynucleotide segments are indicated.

Table 4 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with polynucleotide segments of each template sequence as defined by the indicated "start" and "stop" nucleotide positions. The reading frames of the polynucleotide segments are shown, and the  
25 polypeptides encoded by the polynucleotide segments constitute either signal peptide (SP) or transmembrane (TM) domains, as indicated. The membrane topology of the encoded polypeptide sequence is indicated, the N-terminus (N) listed as being oriented to either the cytosolic (N in) or non-cytosolic (N out) side of the cell membrane or organelle.

Table 5 shows the sequence identification numbers (SEQ ID NO:s) and template  
30 identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with component sequence identification numbers (component IDs) corresponding to each template. The component sequences, which were used to assemble the template sequences, are defined by the indicated "start" and "stop" nucleotide positions along each template.

Table 6 shows the tissue distribution profiles for the templates of the invention.

35 Table 7 shows the sequence identification numbers (SEQ ID NO:s) corresponding to the polypeptides of the present invention, along with the reading frames used to obtain the polypeptide

segments, the lengths of the polypeptide segments, the "start" and "stop" nucleotide positions of the polynucleotide sequences used to define the encoded polypeptide segments, the GenBank hits (GI Numbers), probability scores, and functional annotations corresponding to the GenBank hits.

Table 8 summarizes the bioinformatics tools which are useful for analysis of the polynucleotides of the present invention. The first column of Table 8 lists analytical tools, programs, and algorithms, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences).

## DETAILED DESCRIPTION OF THE INVENTION

Before the nucleic acid sequences and methods are presented, it is to be understood that this invention is not limited to the particular machines, methods, and materials described. Although particular embodiments are described, machines, methods, and materials similar or equivalent to these embodiments may be used to practice the invention. The preferred machines, methods, and materials set forth are not intended to limit the scope of the invention which is limited only by the appended claims.

The singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. All technical and scientific terms have the meanings commonly understood by one of ordinary skill in the art. All publications are incorporated by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are presented and which might be used in connection with the invention. Nothing in the specification is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

### Definitions

As used herein, the lower case "dithp" refers to a nucleic acid sequence, while the upper case "DITHP" refers to an amino acid sequence encoded by dithp. A "full-length" dithp refers to a nucleic acid sequence containing the entire coding region of a gene endogenously expressed in human tissue.

"Adjuvants" are materials such as Freund's adjuvant, mineral gels (aluminum hydroxide), and surface active substances (lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol) which may be administered to increase a host's immunological response.

"Allele" refers to an alternative form of a nucleic acid sequence. Alleles result from a "mutation," a change or an alternative reading of the genetic code. Any given gene may have none, one, or many allelic forms. Mutations which give rise to alleles include deletions, additions, or



substitutions of nucleotides. Each of these changes may occur alone, or in combination with the others, one or more times in a given nucleic acid sequence. The present invention encompasses allelic dithp.

“Amino acid sequence” refers to a peptide, a polypeptide, or a protein of either natural or synthetic origin. The amino acid sequence is not limited to the complete, endogenous amino acid sequence and may be a fragment, epitope, variant, or derivative of a protein expressed by a nucleic acid sequence.

“Amplification” refers to the production of additional copies of a sequence and is carried out using polymerase chain reaction (PCR) technologies well known in the art.

“Antibody” refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind DITHP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

“Antisense sequence” refers to a sequence capable of specifically hybridizing to a target sequence. The antisense sequence may include DNA, RNA, or any nucleic acid mimic or analog such as peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine.

“Antisense sequence” refers to a sequence capable of specifically hybridizing to a target sequence. The antisense sequence can be DNA, RNA, or any nucleic acid mimic or analog.

“Antisense technology” refers to any technology which relies on the specific hybridization of an antisense sequence to a target sequence.

A “bin” is a portion of computer memory space used by a computer program for storage of data, and bounded in such a manner that data stored in a bin may be retrieved by the program.

“Biologically active” refers to an amino acid sequence having a structural, regulatory, or biochemical function of a naturally occurring amino acid sequence.

“Clone joining” is a process for combining gene bins based upon the bins' containing sequence information from the same clone. The sequences may assemble into a primary gene transcript as well as one or more splice variants.

“Complementary” describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing (5'-A-G-T-3' pairs with its complement 3'-T-C-A-5').

A "component sequence" is a nucleic acid sequence selected by a computer program such as PHRED and used to assemble a consensus or template sequence from one or more component sequences.

5 A "consensus sequence" or "template sequence" is a nucleic acid sequence which has been assembled from overlapping sequences, using a computer program for fragment assembly such as the GELVIEW fragment assembly system (Genetics Computer Group (GCG), Madison WI) or using a relational database management system (RDMS).

10 "Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative substitutions.

	Original Residue	Conservative Substitution
15	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
20	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
25	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
30	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

35

Conservative substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

40 "Deletion" refers to a change in either a nucleic or amino acid sequence in which at least one nucleotide or amino acid residue, respectively, is absent.

"Derivative" refers to the chemical modification of a nucleic acid sequence, such as by replacement of hydrogen by an alkyl, acyl, amino, hydroxyl, or other group.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

5       The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

"E-value" refers to the statistical probability that a match between two sequences occurred by chance.

10       "Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

15       A "fragment" is a unique portion of dithp or DITHP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 10 to 1000 contiguous amino acid residues or nucleotides. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous amino acid residues or nucleotides in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing and the figures, may be encompassed by the present embodiments.

20       A fragment of dithp comprises a region of unique polynucleotide sequence that specifically identifies dithp, for example, as distinct from any other sequence in the same genome. A fragment of dithp is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish dithp from related polynucleotide sequences. The precise length of a fragment of dithp and the region of dithp to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

30       A fragment of DITHP is encoded by a fragment of dithp. A fragment of DITHP comprises a region of unique amino acid sequence that specifically identifies DITHP. For example, a fragment of DITHP is useful as an immunogenic peptide for the development of antibodies that specifically recognize DITHP. The precise length of a fragment of DITHP and the region of DITHP to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

35

A "full length" nucleotide sequence is one containing at least a start site for translation to a protein sequence, followed by an open reading frame and a stop site, and encoding a "full length" polypeptide.

"Hit" refers to a sequence whose annotation will be used to describe a given template.

- 5 Criteria for selecting the top hit are as follows: if the template has one or more exact nucleic acid matches, the top hit is the exact match with highest percent identity. If the template has no exact matches but has significant protein hits, the top hit is the protein hit with the lowest E-value. If the template has no significant protein hits, but does have significant non-exact nucleotide hits, the top hit is the nucleotide hit with the lowest E-value.

- 10 "Homology" refers to sequence similarity either between a reference nucleic acid sequence and at least a fragment of a dithp or between a reference amino acid sequence and a fragment of a DITHP.

- "Hybridization" refers to the process by which a strand of nucleotides anneals with a complementary strand through base pairing. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under defined annealing conditions, and remain hybridized after the "washing" step. The defined hybridization conditions include the annealing conditions and the washing step(s), the latter of which is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid probes that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency.
- 15  
20

- Generally, stringency of hybridization is expressed with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization is well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.
- 25  
30

- High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, or 55°C may be used. SSC concentration may be varied from about 0.2 to 2 x SSC, with SDS being present at about 0.1%.
- 35 Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Useful variations on

these conditions will be readily apparent to those skilled in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their resultant proteins.

Other parameters, such as temperature, salt concentration, and detergent concentration may be varied to achieve the desired stringency. Denaturants, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as RNA:DNA hybridizations. Appropriate hybridization conditions are routinely determinable by one of ordinary skill in the art.

"Immunologically active" or "immunogenic" describes the potential for a natural, recombinant, or synthetic peptide, epitope, polypeptide, or protein to induce antibody production in appropriate animals, cells, or cell lines.

"Insertion" or "addition" refers to a change in either a nucleic or amino acid sequence in which at least one nucleotide or residue, respectively, is added to the sequence.

"Labeling" refers to the covalent or noncovalent joining of a polynucleotide, polypeptide, or antibody with a reporter molecule capable of producing a detectable or measurable signal.

"Microarray" is any arrangement of nucleic acids, amino acids, antibodies, etc., on a substrate. The substrate may be a solid support such as beads, glass, paper, nitrocellulose, nylon, or an appropriate membrane.

"Linkers" are short stretches of nucleotide sequence which may be added to a vector or a dithp to create restriction endonuclease sites to facilitate cloning. "Polylinkers" are engineered to incorporate multiple restriction enzyme sites and to provide for the use of enzymes which leave 5' or 3' overhangs (e.g., BamHI, EcoRI, and HindIII) and those which provide blunt ends (e.g., EcoRV, SnaBI, and StuI).

"Naturally occurring" refers to an endogenous polynucleotide or polypeptide that may be isolated from viruses or prokaryotic or eukaryotic cells.

"Nucleic acid sequence" refers to the specific order of nucleotides joined by phosphodiester bonds in a linear, polymeric arrangement. Depending on the number of nucleotides, the nucleic acid sequence can be considered an oligomer, oligonucleotide, or polynucleotide. The nucleic acid can be DNA, RNA, or any nucleic acid analog, such as PNA, may be of genomic or synthetic origin, may be either double-stranded or single-stranded, and can represent either the sense or antisense (complementary) strand.

"Oligomer" refers to a nucleic acid sequence of at least about 6 nucleotides and as many as about 60 nucleotides, preferably about 15 to 40 nucleotides, and most preferably between about 20 and 30 nucleotides, that may be used in hybridization or amplification technologies. Oligomers may be used as, e.g., primers for PCR, and are usually chemically synthesized.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to a DNA mimic in which nucleotide bases are attached to a pseudopeptide backbone to increase stability. PNAs, also designated antigene agents, can prevent gene expression by targeting complementary messenger RNA.

The phrases "percent identity" and "% identity", as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and Sharp, P.M. (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to determine alignment between a known polynucleotide sequence and other sequences on a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2/>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

5 *Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

*Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example,  
 10 as defined by a particular SEQ ID number, or may be measured over a shorter length, for example,  
 over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at  
 least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous  
 nucleotides. Such lengths are exemplary only, and it is understood that any fragment length  
 supported by the sequences shown herein, in figures or Sequence Listings, may be used to describe a  
 15 length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode  
 similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes  
 in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid  
 sequences that all encode substantially the same protein.

20 The phrases "percent identity" and "% identity", as applied to polypeptide sequences, refer to  
 the percentage of residue matches between at least two polypeptide sequences aligned using a  
 standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some  
 alignment methods take into account conservative amino acid substitutions. Such conservative  
 substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity of the  
 25 substituted residue, thus preserving the structure (and therefore function) of the folded polypeptide.

Percent identity between polypeptide sequences may be determined using the default  
 parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e  
 sequence alignment program (described and referenced above). For pairwise alignments of  
 polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap  
 30 penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default  
 residue weight table. As with polynucleotide alignments, the percent identity is reported by  
 CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise  
 comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9  
 35 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Open Gap: 11 and Extension Gap: 1 penalty*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 3*

5 *Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least  
10 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in figures or Sequence Listings, may be used to describe a length over which percentage identity may be measured.

"Post-translational modification" of a DITHP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in  
15 the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu and the DITHP.

"Probe" refers to dithp or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands,  
20 chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

25 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the  
30 specification, including the figures and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A  
35 Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be



derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

"Purified" refers to molecules, either polynucleotides or polypeptides that are isolated or separated from their natural environment and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other compounds with which they are naturally associated.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

“Regulatory element” refers to a nucleic acid sequence from nontranslated regions of a gene, and includes enhancers, promoters, introns, and 3’ untranslated regions, which interact with host proteins to carry out or regulate transcription or translation.

“Reporter” molecules are chemical or biochemical moieties used for labeling a nucleic acid, an amino acid, or an antibody. They include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An “RNA equivalent,” in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

“Sample” is used in its broadest sense. Samples may contain nucleic or amino acids, antibodies, or other materials, and may be derived from any source (e.g., bodily fluids including, but not limited to, saliva, blood, and urine; chromosome(s), organelles, or membranes isolated from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; and cleared cells or tissues or blots or imprints from such cells or tissues).

“Specific binding” or “specifically binding” refers to the interaction between a protein or peptide and its agonist, antibody, antagonist, or other binding partner. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope “A,” the presence of a polypeptide containing epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

“Substitution” refers to the replacement of at least one nucleotide or amino acid by a different nucleotide or amino acid.

“Substrate” refers to any suitable rigid or semi-rigid support including, e.g., membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles or capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A “transcript image” refers to the collective pattern of gene expression by a particular tissue or cell type under given conditions at a given time.

“Transformation” refers to a process by which exogenous DNA enters a recipient cell. Transformation may occur under natural or artificial conditions using various methods well known in

the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed.

“Transformants” include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as cells which transiently express inserted DNA or RNA.

A “transgenic organism,” as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), *supra*.

A “variant” of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 25% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 30%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. The variant may result in “conservative” amino acid changes which do not affect structural and/or chemical properties. A variant may be described as, for example, an “allelic” (as defined above), “splice,” “species,” or “polymorphic” variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass “single nucleotide polymorphisms” (SNPs) in which the polynucleotide sequence varies by one base.

The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

In an alternative, variants of the polynucleotides of the present invention may be generated through recombinant methods. One possible method is a DNA shuffling technique such as

5 MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of DITHP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene

10 variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random

15 point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

20 A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least

25 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater identity over a certain defined length of one of the polypeptides.

## THE INVENTION

In a particular embodiment, cDNA sequences derived from human tissues and cell lines were

30 aligned based on nucleotide sequence identity and assembled into "consensus" or "template" sequences which are designated by the template identification numbers (template IDs) in column 2 of Table 2. The sequence identification numbers (SEQ ID NO:s) corresponding to the template IDs are shown in column 1. The template sequences have similarity to GenBank sequences, or "hits," as designated by the GI Numbers in column 3. The statistical probability of each GenBank hit is

35 indicated by a probability score in column 4, and the functional annotation corresponding to each GenBank hit is listed in column 5.

The invention incorporates the nucleic acid sequences of these templates as disclosed in the Sequence Listing and the use of these sequences in the diagnosis and treatment of disease states characterized by defects in human molecules. The invention further utilizes these sequences in hybridization and amplification technologies, and in particular, in technologies which assess gene expression patterns correlated with specific cells or tissues and their responses *in vivo* or *in vitro* to pharmaceutical agents, toxins, and other treatments. In this manner, the sequences of the present invention are used to develop a transcript image for a particular cell or tissue.

#### Derivation of Nucleic Acid Sequences

cDNA was isolated from libraries constructed using RNA derived from normal and diseased human tissues and cell lines. The human tissues and cell lines used for cDNA library construction were selected from a broad range of sources to provide a diverse population of cDNAs representative of gene transcription throughout the human body. Descriptions of the human tissues and cell lines used for cDNA library construction are provided in the LIFESEQ database (Incyte Genomics, Inc. (Incyte), Palo Alto CA). Human tissues were broadly selected from, for example, cardiovascular, dermatologic, endocrine, gastrointestinal, hematopoietic/immune system, musculoskeletal, neural, reproductive, and urologic sources.

Cell lines used for cDNA library construction were derived from, for example, leukemic cells, teratocarcinomas, neuroepitheliomas, cervical carcinoma, lung fibroblasts, and endothelial cells. Such cell lines include, for example, THP-1, Jurkat, HUVEC, hNT2, WI38, HeLa, and other cell lines commonly used and available from public depositories (American Type Culture Collection, Manassas VA). Prior to mRNA isolation, cell lines were untreated, treated with a pharmaceutical agent such as 5'-aza-2'-deoxycytidine, treated with an activating agent such as lipopolysaccharide in the case of leukocytic cell lines, or, in the case of endothelial cell lines, subjected to shear stress.

#### Sequencing of the cDNAs

Methods for DNA sequencing are well known in the art. Conventional enzymatic methods employ the Klenow fragment of DNA polymerase I, SEQUENASE DNA polymerase (U.S. Biochemical Corporation, Cleveland OH), Taq polymerase (Applied Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Inc. (Amersham Pharmacia Biotech), Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies Inc. (Life Technologies), Gaithersburg MD), to extend the nucleic acid sequence from an oligonucleotide primer annealed to the DNA template of interest. Methods have been developed for the use of both single-stranded and double-stranded templates. Chain termination reaction products may be electrophoresed on urea-polyacrylamide gels and detected either by autoradiography (for radioisotope-labeled nucleotides) or

by fluorescence (for fluorophore-labeled nucleotides). Automated methods for mechanized reaction preparation, sequencing, and analysis using fluorescence detection methods have been developed. Machines used to prepare cDNAs for sequencing can include the MICROLAB 2200 liquid transfer system (Hamilton Company (Hamilton), Reno NV), Peltier thermal cycler (PTC200; MJ Research, Inc. (MJ Research), Watertown MA), and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing can be carried out using, for example, the ABI 373 or 377 (Applied Biosystems) or MEGABACE 1000 (Molecular Dynamics, Inc. (Molecular Dynamics), Sunnyvale CA) DNA sequencing systems, or other automated and manual sequencing systems well known in the art.

The nucleotide sequences of the Sequence Listing have been prepared by current, state-of-the-art, automated methods and, as such, may contain occasional sequencing errors or unidentified nucleotides. Such unidentified nucleotides are designated by an N. These infrequent unidentified bases do not represent a hindrance to practicing the invention for those skilled in the art. Several methods employing standard recombinant techniques may be used to correct errors and complete the missing sequence information. (See, e.g., those described in Ausubel, F.M. et al. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY; and Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.)

#### Assembly of cDNA Sequences

Human polynucleotide sequences may be assembled using programs or algorithms well known in the art. Sequences to be assembled are related, wholly or in part, and may be derived from a single or many different transcripts. Assembly of the sequences can be performed using such programs as PHRAP (Phils Revised Assembly Program) and the GELVIEW fragment assembly system (GCG), or other methods known in the art.

Alternatively, cDNA sequences are used as "component" sequences that are assembled into "template" or "consensus" sequences as follows. Sequence chromatograms are processed, verified, and quality scores are obtained using PHRED. Raw sequences are edited using an editing pathway known as Block 1 (See, e.g., the LIFESEQ Assembled User Guide, Incyte Genomics, Palo Alto, CA). A series of BLAST comparisons is performed and low-information segments and repetitive elements (e.g., dinucleotide repeats, Alu repeats, etc.) are replaced by "n's", or masked, to prevent spurious matches. Mitochondrial and ribosomal RNA sequences are also removed. The processed sequences are then loaded into a relational database management system (RDMS) which assigns edited sequences to existing templates, if available. When additional sequences are added into the RDMS, a process is initiated which modifies existing templates or creates new templates from works in progress (i.e., nonfinal assembled sequences) containing queued sequences or the sequences

themselves. After the new sequences have been assigned to templates, the templates can be merged into bins. If multiple templates exist in one bin, the bin can be split and the templates reannotated.

Once gene bins have been generated based upon sequence alignments, bins are "clone joined" based upon clone information. Clone joining occurs when the 5' sequence of one clone is present in one bin and the 3' sequence from the same clone is present in a different bin, indicating that the two bins should be merged into a single bin. Only bins which share at least two different clones are merged.

A resultant template sequence may contain either a partial or a full length open reading frame, or all or part of a genetic regulatory element. This variation is due in part to the fact that the full length cDNAs of many genes are several hundred, and sometimes several thousand, bases in length. With current technology, cDNAs comprising the coding regions of large genes cannot be cloned because of vector limitations, incomplete reverse transcription of the mRNA, or incomplete "second strand" synthesis. Template sequences may be extended to include additional contiguous sequences derived from the parent RNA transcript using a variety of methods known to those of skill in the art. Extension may thus be used to achieve the full length coding sequence of a gene.

#### Analysis of the cDNA Sequences

The cDNA sequences are analyzed using a variety of programs and algorithms which are well known in the art. (See, e.g., Ausubel, 1997, supra, Chapter 7.7; Meyers, R.A. (Ed.) (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853; and Table 8.) These analyses comprise both reading frame determinations, e.g., based on triplet codon periodicity for particular organisms (Fickett, J.W. (1982) *Nucleic Acids Res.* 10:5303-5318); analyses of potential start and stop codons; and homology searches.

Computer programs known to those of skill in the art for performing computer-assisted searches for amino acid and nucleic acid sequence similarity, include, for example, Basic Local Alignment Search Tool (BLAST; Altschul, S.F. (1993) *J. Mol. Evol.* 36:290-300; Altschul, S.F. et al. (1990) *J. Mol. Biol.* 215:403-410). BLAST is especially useful in determining exact matches and comparing two sequence fragments of arbitrary but equal lengths, whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user (Karlin, S. et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:841-845). Using an appropriate search tool (e.g., BLAST or HMM), GenBank, SwissProt, BLOCKS, PFAM and other databases may be searched for sequences containing regions of homology to a query dithp or DITHP of the present invention.

Other approaches to the identification, assembly, storage, and display of nucleotide and polypeptide sequences are provided in "Relational Database for Storing Biomolecule Information," U.S.S.N. 08/947,845, filed October 9, 1997; "Project-Based Full-Length Biomolecular Sequence

Database," U.S.S.N. 08/811,758, filed March 6, 1997; and "Relational Database and System for Storing Information Relating to Biomolecular Sequences," U.S.S.N. 09/034,807, filed March 4, 1998, all of which are incorporated by reference herein in their entirety.

Protein hierarchies can be assigned to the putative encoded polypeptide based on, e.g., motif, BLAST, or biological analysis. Methods for assigning these hierarchies are described, for example, in "Database System Employing Protein Function Hierarchies for Viewing Biomolecular Sequence Data," U.S.S.N. 08/812,290, filed March 6, 1997, incorporated herein by reference.

#### Identification of Human Diagnostic and Therapeutic Molecules Encoded by dithp

The identities of the DITHP encoded by the dithp of the present invention were obtained by analysis of the assembled cDNA sequences.

SEQ ID NO:276, SEQ ID NO:277, SEQ ID NO:278, SEQ ID NO:279, SEQ ID NO:280, SEQ ID NO:281, SEQ ID NO:282, SEQ ID NO:283, SEQ ID NO:284, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, and SEQ ID NO:291, encoded by SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, and SEQ ID NO:16, respectively, are, for example, human enzyme molecules.

SEQ ID NO:292, SEQ ID NO:293, SEQ ID NO:294, SEQ ID NO:295, and SEQ ID NO:296, encoded by SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, and SEQ ID NO:21, respectively, are, for example, extracellular information transmission molecules.

SEQ ID NO:297, SEQ ID NO:298, SEQ ID NO:299, SEQ ID NO:300, SEQ ID NO:301, and SEQ ID NO:302, encoded by SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, and SEQ ID NO:27, respectively, are, for example, receptor molecules.

SEQ ID NO:303, SEQ ID NO:304, SEQ ID NO:305, SEQ ID NO:306, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:309, SEQ ID NO:310, SEQ ID NO:311, SEQ ID NO:312, SEQ ID NO:313, SEQ ID NO:314, SEQ ID NO:315, SEQ ID NO:316, SEQ ID NO:317, and SEQ ID NO:318, encoded by SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43, respectively, are, for example, intracellular signaling molecules.

SEQ ID NO:319, SEQ ID NO:320, SEQ ID NO:321, SEQ ID NO:322, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:325, SEQ ID NO:326, SEQ ID NO:327, SEQ ID NO:328, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:331, SEQ ID NO:332, SEQ ID NO:333, SEQ ID NO:334, SEQ ID NO:335, SEQ ID NO:336, SEQ ID NO:337, SEQ ID NO:338, SEQ ID NO:339, SEQ ID NO:340, SEQ ID NO:341, SEQ ID NO:342, SEQ ID NO:343, SEQ ID NO:344, SEQ ID NO:345, SEQ ID



NO:346, SEQ ID NO:347, SEQ ID NO:348, SEQ ID NO:349, SEQ ID NO:350, SEQ ID NO:351, SEQ ID NO:352, SEQ ID NO:353, SEQ ID NO:354, SEQ ID NO:355, SEQ ID NO:356, SEQ ID NO:357, SEQ ID NO:358, SEQ ID NO:359, SEQ ID NO:360, SEQ ID NO:361, SEQ ID NO:362, SEQ ID NO:363, SEQ ID NO:364, SEQ ID NO:365, SEQ ID NO:366, SEQ ID NO:367, SEQ ID NO:368, SEQ ID NO:369, SEQ ID NO:370, SEQ ID NO:371, SEQ ID NO:372, SEQ ID NO:373, SEQ ID NO:374, SEQ ID NO:375, SEQ ID NO:376, SEQ ID NO:377, SEQ ID NO:378, SEQ ID NO:379, SEQ ID NO:380, SEQ ID NO:381, SEQ ID NO:382, SEQ ID NO:383, SEQ ID NO:384, SEQ ID NO:385, SEQ ID NO:386, SEQ ID NO:387, SEQ ID NO:388, SEQ ID NO:389, SEQ ID NO:390, SEQ ID NO:391, SEQ ID NO:392, SEQ ID NO:393, SEQ ID NO:394, SEQ ID NO:395, SEQ ID NO:396, and SEQ ID NO:397, encoded by SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, and SEQ ID NO:121, respectively, are, for example, transcription factor molecules.

SEQ ID NO:398, SEQ ID NO:399, SEQ ID NO:400, SEQ ID NO:401, SEQ ID NO:402, SEQ ID NO:403, SEQ ID NO:404, and SEQ ID NO:405, encoded by SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, and SEQ ID NO:129, respectively, are, for example, membrane transport molecules.

SEQ ID NO:406, SEQ ID NO:407, SEQ ID NO:408, SEQ ID NO:409, SEQ ID NO:410, SEQ ID NO:411, SEQ ID NO:412, SEQ ID NO:413, SEQ ID NO:414, SEQ ID NO:415, SEQ ID NO:416, SEQ ID NO:417, SEQ ID NO:418, SEQ ID NO:419, SEQ ID NO:420, SEQ ID NO:421, SEQ ID NO:422, and SEQ ID NO:423, encoded by SEQ ID NO:130, SEQ ID NO:131, SEQ ID NO:132, SEQ ID NO:133, SEQ ID NO:134, SEQ ID NO:135, SEQ ID NO:136, SEQ ID NO:137, SEQ ID NO:138, SEQ ID NO:139, SEQ ID NO:140, SEQ ID NO:141, SEQ ID NO:142, SEQ ID NO:143, SEQ ID NO:144, SEQ ID NO:145, SEQ ID NO:146, and SEQ ID NO:147, respectively, are, for example, protein modification and maintenance molecules.

SEQ ID NO:424, SEQ ID NO:425, SEQ ID NO:426, SEQ ID NO:427, SEQ ID NO:428, SEQ ID NO:429, SEQ ID NO:430, SEQ ID NO:431, SEQ ID NO:432, SEQ ID NO:433, SEQ ID NO:434, and SEQ ID NO:435, encoded by SEQ ID NO:148, SEQ ID NO:149, SEQ ID NO:150, SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153, SEQ ID NO:154, SEQ ID NO:155, SEQ ID NO:156, 5 SEQ ID NO:157, SEQ ID NO:158, and SEQ ID NO:159, respectively, are, for example, nucleic acid synthesis and modification molecules.

SEQ ID NO:436, encoded by SEQ ID NO:160 is, for example, an adhesion molecule.

SEQ ID NO:437, SEQ ID NO:438, and SEQ ID NO:439, encoded by SEQ ID NO:161, SEQ ID NO:162, and SEQ ID NO:163, respectively, are, for example, antigen recognition molecules.

10 SEQ ID NO:440, SEQ ID NO:441, SEQ ID NO:442, and SEQ ID NO:443, encoded by SEQ ID NO:164, SEQ ID NO:165, SEQ ID NO:166, and SEQ ID NO:167, respectively, are, for example, electron transfer associated molecules.

SEQ ID NO:444, SEQ ID NO:445, SEQ ID NO:446, SEQ ID NO:447, SEQ ID NO:448, and SEQ ID NO:449, encoded by SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, SEQ ID NO:171, 15 SEQ ID NO:172, and SEQ ID NO:173, respectively, are, for example, secreted/extracellular matrix molecules.

SEQ ID NO:450, SEQ ID NO:451, SEQ ID NO:452, SEQ ID NO:453, SEQ ID NO:454, SEQ ID NO:455, SEQ ID NO:456, SEQ ID NO:457, SEQ ID NO:458, SEQ ID NO:459, SEQ ID NO:460, and SEQ ID NO:461, encoded by SEQ ID NO:174, SEQ ID NO:175, SEQ ID NO:176, SEQ ID NO:177, SEQ ID NO:178, SEQ ID NO:179, SEQ ID NO:180, SEQ ID NO:181, SEQ ID NO:182, 20 SEQ ID NO:183, SEQ ID NO:184, and SEQ ID NO:185, respectively, are, for example, cytoskeletal molecules.

SEQ ID NO:462, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:465, SEQ ID NO:466, SEQ ID NO:467, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, and SEQ ID NO:471, encoded by 25 SEQ ID NO:186, SEQ ID NO:187, SEQ ID NO:188, SEQ ID NO:189, SEQ ID NO:190, SEQ ID NO:191, SEQ ID NO:192, SEQ ID NO:193, SEQ ID NO:194, and SEQ ID NO:195, respectively, are, for example, cell membrane molecules.

SEQ ID NO:472, SEQ ID NO:473, SEQ ID NO:474, SEQ ID NO:475, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:478, SEQ ID NO:479, SEQ ID NO:480, SEQ ID NO:481, SEQ ID NO:482, 30 SEQ ID NO:483, SEQ ID NO:484, SEQ ID NO:485, SEQ ID NO:486, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:492, SEQ ID NO:493, SEQ ID NO:494, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:498, SEQ ID NO:499, SEQ ID NO:500, SEQ ID NO:501, SEQ ID NO:502, SEQ ID NO:503, SEQ ID NO:504, SEQ ID NO:505, SEQ ID NO:506, SEQ ID NO:507, SEQ ID NO:508, SEQ ID NO:509, SEQ ID NO:510, SEQ ID NO:511, SEQ ID NO:512, SEQ ID NO:513, SEQ ID NO:514, SEQ ID NO:515, 35 SEQ ID NO:516, SEQ ID NO:517, and SEQ ID NO:518, encoded by SEQ ID NO:196, SEQ ID

NO:197, SEQ ID NO:198, SEQ ID NO:199, SEQ ID NO:200, SEQ ID NO:201, SEQ ID NO:202, SEQ ID NO:203, SEQ ID NO:204, SEQ ID NO:205, SEQ ID NO:206, SEQ ID NO:207, SEQ ID NO:208, SEQ ID NO:209, SEQ ID NO:210, SEQ ID NO:211, SEQ ID NO:212, SEQ ID NO:213, SEQ ID NO:214, SEQ ID NO:215, SEQ ID NO:216, SEQ ID NO:217, SEQ ID NO:218, SEQ ID NO:219, SEQ ID NO:220, SEQ ID NO:221, SEQ ID NO:222, SEQ ID NO:223, SEQ ID NO:224, SEQ ID NO:225, SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, SEQ ID NO:229, SEQ ID NO:230, SEQ ID NO:231, SEQ ID NO:231, SEQ ID NO:232, SEQ ID NO:233, SEQ ID NO:234, SEQ ID NO:235, SEQ ID NO:236, SEQ ID NO:237, SEQ ID NO:238, SEQ ID NO:239, SEQ ID NO:240, and SEQ ID NO:241, respectively, are, for example, ribosomal molecules.

SEQ ID NO:519, SEQ ID NO:520, and SEQ ID NO:521, encoded by SEQ ID NO:242, SEQ ID NO:243, and SEQ ID NO:244, respectively, are, for example, chromatin molecules.

SEQ ID NO:522, SEQ ID NO:523, SEQ ID NO:524, SEQ ID NO:525, SEQ ID NO:526, SEQ ID NO:527, SEQ ID NO:528, SEQ ID NO:529, and SEQ ID NO:530, encoded by SEQ ID NO:245, SEQ ID NO:246, SEQ ID NO:247, SEQ ID NO:248, SEQ ID NO:249, SEQ ID NO:250, SEQ ID NO:251, SEQ ID NO:252, and SEQ ID NO:253, respectively, are, for example, organelle associated molecules.

SEQ ID NO:531, SEQ ID NO:532, SEQ ID NO:533, SEQ ID NO:534, SEQ ID NO:535, SEQ ID NO:536, SEQ ID NO:537, SEQ ID NO:538, SEQ ID NO:539, SEQ ID NO:540, SEQ ID NO:541, SEQ ID NO:542, SEQ ID NO:543, and SEQ ID NO:544, encoded by SEQ ID NO:254, SEQ ID NO:255, SEQ ID NO:256, SEQ ID NO:256, SEQ ID NO:257, SEQ ID NO:258, SEQ ID NO:259, SEQ ID NO:260, SEQ ID NO:261, SEQ ID NO:262, SEQ ID NO:263, SEQ ID NO:264, SEQ ID NO:265, and SEQ ID NO:266, respectively, are, for example, biochemical pathway molecules.

SEQ ID NO:545, SEQ ID NO:546, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:552, and SEQ ID NO:553, encoded by SEQ ID NO:267, SEQ ID NO:268, SEQ ID NO:269, SEQ ID NO:270, SEQ ID NO:271, SEQ ID NO:272, SEQ ID NO:273, SEQ ID NO:274, and SEQ ID NO:275, respectively, are, for example, molecules associated with growth and development.

#### Sequences of Human Diagnostic and Therapeutic Molecules

The dithp of the present invention may be used for a variety of diagnostic and therapeutic purposes. For example, a dithp may be used to diagnose a particular condition, disease, or disorder associated with human molecules. Such conditions, diseases, and disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma,

and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder, such as inflammation, actinic keratosis, acquired

5 immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis,

10 Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma,

15 Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus,

20 flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio,

25 brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malassezia, histoplasma, or other mycosis-causing fungal agent; and an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal

30 protozoa such as giardia, trichomonas, tissue nematode such as trichinella, intestinal nematode such as ascaris, lymphatic filarial nematode, trematode such as schistosoma, and cestode such as tapeworm; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation),

35 Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and

neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; an endocrine disorder such as a disorder of the hypothalamus and/or pituitary resulting from lesions such as a primary brain tumor, adenoma, infarction associated with pregnancy, hypophysectomy, aneurysm, vascular malformation, thrombosis, infection, immunological disorder, and complication due to head trauma; a disorder associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism; a disorder associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma; a disorder associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism; a disorder associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease; a disorder associated with hyperparathyroidism including Conn disease (chronic hypercalcemia); a pancreatic disorder such as Type I or Type II diabetes mellitus and associated complications; a disorder associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease; a disorder associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbation of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, a hypergonadal disorder associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5  $\alpha$ -reductase, and gynecomastia; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, pentosuria phenylketonuria, pseudovitamin D-deficiency rickets; disorders of carbohydrate metabolism such as congenital type II dyserythropoietic anemia, diabetes, insulin-dependent diabetes mellitus, non-insulin-dependent diabetes mellitus, fructose-1,6-diphosphatase deficiency, galactosemia, glucagonoma, hereditary fructose intolerance,

hypoglycemia, mannosidosis, neuraminidase deficiency, obesity, galactose epimerase deficiency, glycogen storage diseases, lysosomal storage diseases, fructosuria, pentosuria, and inherited abnormalities of pyruvate metabolism; disorders of lipid metabolism such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency,

5 myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such as Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM<sub>2</sub> gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipid adrenal hyperplasia, minimal change disease,

10 lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity; and disorders of copper metabolism such as Menke's disease, Wilson's disease, and Ehlers-Danlos

15 syndrome type IX; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural

20 empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis; encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central

25 nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorder, a peripheral nervous system disorder, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathy, myasthenia gravis, periodic paralysis, a mental disorder including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia,

30 catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a gastrointestinal disorder including ulcerative colitis, gastric and duodenal ulcers, cystinuria, dibasicaminoaciduria, hypercystinuria, lysinuria, hartnup disease, tryptophan malabsorption, methionine malabsorption, histidinuria, iminoglycinuria, dicarboxylicaminoaciduria, cystinosis, renal glycosuria, hypouricemia, familial hypophosphatemic

35 rickets, congenital chloridorrhea, distal renal tubular acidosis, Menkes' disease, Wilson's disease, lethal diarrhea, juvenile pernicious anemia, folate malabsorption, adrenoleukodystrophy, hereditary

myoglobinuria, and Zellweger syndrome; a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, and polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, glucose-galactose malabsorption syndrome, hypercholesterolemia, Cushing's disease, and Addison's disease; and a connective tissue disorder such as osteogenesis imperfecta, Ehlers-Danlos syndrome, chondrodysplasias, Marfan syndrome, Alport syndrome, familial aortic aneurysm, achondroplasia, mucopolysaccharidoses, osteoporosis, osteopetrosis, Paget's disease, rickets, osteomalacia, hyperparathyroidism, renal osteodystrophy, osteonecrosis, osteomyelitis, osteoma, osteoid osteoma, osteoblastoma, osteosarcoma, osteochondroma, chondroma, chondroblastoma, chondromyxoid fibroma, chondrosarcoma, fibrous cortical defect, nonossifying fibroma, fibrous dysplasia, fibrosarcoma, malignant fibrous histiocytoma, Ewing's sarcoma, primitive neuroectodermal tumor, giant cell tumor, osteoarthritis, rheumatoid arthritis, ankylosing spondyloarthritis, Reiter's syndrome, psoriatic arthritis, enteropathic arthritis, infectious arthritis, gout, gouty arthritis, calcium pyrophosphate crystal deposition disease, ganglion, synovial cyst, villonodular synovitis, systemic sclerosis, Dupuytren's contracture, hepatic fibrosis, lupus erythematosus, mixed connective tissue disease, epidermolysis bullosa simplex, bullous congenital ichthyosiform erythroderma (epidermolytic hyperkeratosis), non-epidermolytic and epidermolytic palmoplantar keratoderma, ichthyosis bullosa of Siemens, pachyonychia congenita, and white sponge nevus. The dithp can be used to detect the presence of, or to quantify the amount of, a dithp-related polynucleotide in a sample. This information is then compared to information obtained from appropriate reference samples, and a diagnosis is established. Alternatively, a polynucleotide complementary to a given dithp can inhibit or inactivate a therapeutically relevant gene related to the dithp.

35

#### Analysis of dithp Expression Patterns

The expression of dithp may be routinely assessed by hybridization-based methods to determine, for example, the tissue-specificity, disease-specificity, or developmental stage-specificity of dithp expression. For example, the level of expression of dithp may be compared among different cell types or tissues, among diseased and normal cell types or tissues, among cell types or tissues at  
5 different developmental stages, or among cell types or tissues undergoing various treatments. This type of analysis is useful, for example, to assess the relative levels of dithp expression in fully or partially differentiated cells or tissues, to determine if changes in dithp expression levels are correlated with the development or progression of specific disease states, and to assess the response of a cell or tissue to a specific therapy, for example, in pharmacological or toxicological studies.  
10 Methods for the analysis of dithp expression are based on hybridization and amplification technologies and include membrane-based procedures such as northern blot analysis, high-throughput procedures that utilize, for example, microarrays, and PCR-based procedures.

#### Hybridization and Genetic Analysis

15 The dithp, their fragments, or complementary sequences, may be used to identify the presence of and/or to determine the degree of similarity between two (or more) nucleic acid sequences. The dithp may be hybridized to naturally occurring or recombinant nucleic acid sequences under appropriately selected temperatures and salt concentrations. Hybridization with a probe based on the nucleic acid sequence of at least one of the dithp allows for the detection of nucleic acid sequences,  
20 including genomic sequences, which are identical or related to the dithp of the Sequence Listing. Probes may be selected from non-conserved or unique regions of at least one of the polynucleotides of SEQ ID NO:1-275 and tested for their ability to identify or amplify the target nucleic acid sequence using standard protocols.

Polynucleotide sequences that are capable of hybridizing, in particular, to those shown in  
25 SEQ ID NO:1-275 and fragments thereof, can be identified using various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions are discussed in "Definitions."

A probe for use in Southern or northern hybridization may be derived from a fragment of a dithp sequence, or its complement, that is up to several hundred nucleotides in length and is either  
30 single-stranded or double-stranded. Such probes may be hybridized in solution to biological materials such as plasmids, bacterial, yeast, or human artificial chromosomes, cleared or sectioned tissues, or to artificial substrates containing dithp. Microarrays are particularly suitable for identifying the presence of and detecting the level of expression for multiple genes of interest by examining gene expression correlated with, e.g., various stages of development; treatment with a drug or compound,  
35 or disease progression. An array analogous to a dot or slot blot may be used to arrange and link polynucleotides to the surface of a substrate using one or more of the following: mechanical



(vacuum), chemical, thermal, or UV bonding procedures. Such an array may contain any number of dithp and may be produced by hand or by using available devices, materials, and machines.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

Probes may be labeled by either PCR or enzymatic techniques using a variety of commercially available reporter molecules. For example, commercial kits are available for radioactive and chemiluminescent labeling (Amersham Pharmacia Biotech) and for alkaline phosphatase labeling (Life Technologies). Alternatively, dithp may be cloned into commercially available vectors for the production of RNA probes. Such probes may be transcribed in the presence of at least one labeled nucleotide (e.g., <sup>32</sup>P-ATP, Amersham Pharmacia Biotech).

Additionally the polynucleotides of SEQ ID NO:1-275 or suitable fragments thereof can be used to isolate full length cDNA sequences utilizing hybridization and/or amplification procedures well known in the art, e.g., cDNA library screening, PCR amplification, etc. The molecular cloning of such full length cDNA sequences may employ the method of cDNA library screening with probes using the hybridization, stringency, washing, and probing strategies described above and in Ausubel, supra, Chapters 3, 5, and 6. These procedures may also be employed with genomic libraries to isolate genomic sequences of dithp in order to analyze, e.g., regulatory elements.

### Genetic Mapping

Gene identification and mapping are important in the investigation and treatment of almost all conditions, diseases, and disorders. Cancer, cardiovascular disease, Alzheimer's disease, arthritis, diabetes, and mental illnesses are of particular interest. Each of these conditions is more complex than the single gene defects of sickle cell anemia or cystic fibrosis, with select groups of genes being predictive of predisposition for a particular condition, disease, or disorder. For example, cardiovascular disease may result from malfunctioning receptor molecules that fail to clear cholesterol from the bloodstream, and diabetes may result when a particular individual's immune system is activated by an infection and attacks the insulin-producing cells of the pancreas. In some studies, Alzheimer's disease has been linked to a gene on chromosome 21; other studies predict a different gene and location. Mapping of disease genes is a complex and reiterative process and generally proceeds from genetic linkage analysis to physical mapping.

As a condition is noted among members of a family, a genetic linkage map traces parts of chromosomes that are inherited in the same pattern as the condition. Statistics link the inheritance of particular conditions to particular regions of chromosomes, as defined by RFLP or other markers.

(See, for example, Lander, E. S. and Botstein, D. (1986) *Proc. Natl. Acad. Sci. USA* 83:7353-7357.) Occasionally, genetic markers and their locations are known from previous studies. More often, however, the markers are simply stretches of DNA that differ among individuals. Examples of genetic linkage maps can be found in various scientific journals or at the Online Mendelian

5 Inheritance in Man (OMIM) World Wide Web site.

In another embodiment of the invention, dithp sequences may be used to generate hybridization probes useful in chromosomal mapping of naturally occurring genomic sequences. Either coding or noncoding sequences of dithp may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a dithp coding  
10 sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J.  
15 et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Meyers, supra, pp. 965-968.) Correlation between the location of dithp on a physical chromosomal map and a specific disorder, or a  
20 predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The dithp sequences may also be used to detect polymorphisms that are genetically linked to the inheritance of a particular condition, disease, or disorder.

In situ hybridization of chromosomal preparations and genetic mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending existing genetic  
25 maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of the corresponding human chromosome is not known. These new marker sequences can be mapped to human chromosomes and may provide valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome has been crudely correlated  
30 by genetic linkage with a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequences of the subject invention may also be used to detect differences in chromosomal architecture due to translocation, inversion, etc., among normal, carrier, or affected individuals.

35 Once a disease-associated gene is mapped to a chromosomal region, the gene must be cloned in order to identify mutations or other alterations (e.g., translocations or inversions) that may be

correlated with disease. This process requires a physical map of the chromosomal region containing the disease-gene of interest along with associated markers. A physical map is necessary for determining the nucleotide sequence of and order of marker genes on a particular chromosomal region. Physical mapping techniques are well known in the art and require the generation of overlapping sets of cloned DNA fragments from a particular organelle, chromosome, or genome. These clones are analyzed to reconstruct and catalog their order. Once the position of a marker is determined, the DNA from that region is obtained by consulting the catalog and selecting clones from that region. The gene of interest is located through positional cloning techniques using hybridization or similar methods.

#### Diagnostic Uses

The dithp of the present invention may be used to design probes useful in diagnostic assays. Such assays, well known to those skilled in the art, may be used to detect or confirm conditions, disorders, or diseases associated with abnormal levels of dithp expression. Labeled probes developed from dithp sequences are added to a sample under hybridizing conditions of desired stringency. In some instances, dithp, or fragments or oligonucleotides derived from dithp, may be used as primers in amplification steps prior to hybridization. The amount of hybridization complex formed is quantified and compared with standards for that cell or tissue. If dithp expression varies significantly from the standard, the assay indicates the presence of the condition, disorder, or disease. Qualitative or quantitative diagnostic methods may include northern, dot blot, or other membrane or dip-stick based technologies or multiple-sample format technologies such as PCR, enzyme-linked immunosorbent assay (ELISA)-like, pin, or chip-based assays.

The probes described above may also be used to monitor the progress of conditions, disorders, or diseases associated with abnormal levels of dithp expression, or to evaluate the efficacy of a particular therapeutic treatment. The candidate probe may be identified from the dithp that are specific to a given human tissue and have not been observed in GenBank or other genome databases. Such a probe may be used in animal studies, preclinical tests, clinical trials, or in monitoring the treatment of an individual patient. In a typical process, standard expression is established by methods well known in the art for use as a basis of comparison, samples from patients affected by the disorder or disease are combined with the probe to evaluate any deviation from the standard profile, and a therapeutic agent is administered and effects are monitored to generate a treatment profile. Efficacy is evaluated by determining whether the expression progresses toward or returns to the standard normal pattern. Treatment profiles may be generated over a period of several days or several months. Statistical methods well known to those skilled in the art may be used to determine the significance of such therapeutic agents.

The polynucleotides are also useful for identifying individuals from minute biological

samples, for example, by matching the RFLP pattern of a sample's DNA to that of an individual's DNA. The polynucleotides of the present invention can also be used to determine the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be  
5 sequenced. Using this technique, an individual can be identified through a unique set of DNA sequences. Once a unique ID database is established for an individual, positive identification of that individual can be made from extremely small tissue samples.

In a particular aspect, oligonucleotide primers derived from the dithp of the invention may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and  
10 deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from dithp are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the  
15 DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms  
20 by comparing the sequences of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc.,  
25 San Diego CA).

DNA-based identification techniques are critical in forensic technology. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using, e.g., PCR, to identify individuals. (See, e.g., Erlich, H. (1992) PCR Technology, Freeman and Co., New York, NY). Similarly, polynucleotides of the  
30 present invention can be used as polymorphic markers.

There is also a need for reagents capable of identifying the source of a particular tissue. Appropriate reagents can comprise, for example, DNA probes or primers prepared from the sequences of the present invention that are specific for particular tissues. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to  
35 screen tissue cultures for contamination.

The polynucleotides of the present invention can also be used as molecular weight markers on

nucleic acid gels or Southern blots, as diagnostic probes for the presence of a specific mRNA in a particular cell type, in the creation of subtracted cDNA libraries which aid in the discovery of novel polynucleotides, in selection and synthesis of oligomers for attachment to an array or other support, and as an antigen to elicit an immune response.

5

#### Disease Model Systems Using dithp

The dithp of the invention or their mammalian homologs may be “knocked out” in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g.,  
10 U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous  
15 recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and  
20 the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

The dithp of the invention may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate  
25 into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

The dithp of the invention can also be used to create “knockin” humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of dithp is injected into animal ES cells, and the injected sequence integrates into the animal cell  
30 genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress dithp, resulting, e.g., in the secretion of DITHP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

35

#### Screening Assays

DITHP encoded by polynucleotides of the present invention may be used to screen for molecules that bind to or are bound by the encoded polypeptides. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the bound molecule. Examples of such molecules include antibodies,  
5 oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a ligand or fragment thereof, a natural substrate, or a structural or functional mimetic. (See, Coligan et al., (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or to at least a fragment of the receptor,  
10 e.g., the active site. In either case, the molecule can be rationally designed using known techniques. Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing the polypeptide or cell membrane fractions which contain the expressed polypeptide are then contacted with a test compound and binding,  
15 stimulation, or inhibition of activity of either the polypeptide or the molecule is analyzed.

An assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. Alternatively, the assay may assess binding in the presence of a labeled competitor.

Additionally, the assay can be carried out using cell-free preparations, polypeptide/molecule  
20 affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay using, e.g., a monoclonal or polyclonal antibody, can measure  
25 polypeptide level in a sample. The antibody can measure polypeptide level by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of the above assays can be used in a diagnostic or prognostic context. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the  
30 assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

#### Transcript Imaging and Toxicological Testing

Another embodiment relates to the use of dithp to develop a transcript image of a tissue or  
35 cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes

and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity pertaining to human molecules for diagnostics and therapeutics.

Transcript images which profile dithp expression may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect dithp expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile dithp expression may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E. F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and Anderson, N.L. (2000) Toxicol. Lett. 112-113:467-71, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with

levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of DITHP encoded by polynucleotides of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for DITHP to quantify the levels of DITHP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-11; Mendoz, L.G. et al. (1999) Biotechniques 27:778-88). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and Seilhamer, J. (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which



alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the DITHP encoded by polynucleotides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the DITHP encoded by polynucleotides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Transcript images may be used to profile dithp expression in distinct tissue types. This process can be used to determine human molecule activity in a particular tissue type relative to this activity in a different tissue type. Transcript images may be used to generate a profile of dithp expression characteristic of diseased tissue. Transcript images of tissues before and after treatment may be used for diagnostic purposes, to monitor the progression of disease, and to monitor the efficacy of drug treatments for diseases which affect the activity of human molecules.

Transcript images of cell lines can be used to assess human molecule activity and/or to identify cell lines that lack or misregulate this activity. Such cell lines may then be treated with pharmaceutical agents, and a transcript image following treatment may indicate the efficacy of these agents in restoring desired levels of this activity. A similar approach may be used to assess the toxicity of pharmaceutical agents as reflected by undesirable changes in human molecule activity. Candidate pharmaceutical agents may be evaluated by comparing their associated transcript images with those of pharmaceutical agents of known effectiveness.

#### Antisense Molecules

The polynucleotides of the present invention are useful in antisense technology. Antisense technology or therapy relies on the modulation of expression of a target protein through the specific binding of an antisense sequence to a target sequence encoding the target protein or directing its

expression. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ; Alama, A. et al. (1997) *Pharmacol. Res.* 36(3):171-178; Crooke, S.T. (1997) *Adv. Pharmacol.* 40:1-49; Sharma, H.W. and R. Narayanan (1995) *Bioessays* 17(12):1055-1063; and Lavrosky, Y. et al. (1997) *Biochem. Mol. Med.* 62(1):11-22.) An antisense sequence is a polynucleotide sequence  
5 capable of specifically hybridizing to at least a portion of the target sequence. Antisense sequences bind to cellular mRNA and/or genomic DNA, affecting translation and/or transcription. Antisense sequences can be DNA, RNA, or nucleic acid mimics and analogs. (See, e.g., Rossi, J.J. et al. (1991) *Antisense Res. Dev.* 1(3):285-288; Lee, R. et al. (1998) *Biochemistry* 37(3):900-1010; Pardridge, W.M. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92(12):5592-5596; and Nielsen, P. E. and Haaima, G.  
10 (1997) *Chem. Soc. Rev.* 96:73-78.) Typically, the binding which results in modulation of expression occurs through hybridization or binding of complementary base pairs. Antisense sequences can also bind to DNA duplexes through specific interactions in the major groove of the double helix.

The polynucleotides of the present invention and fragments thereof can be used as antisense sequences to modify the expression of the polypeptide encoded by dithp. The antisense sequences  
15 can be produced ex vivo, such as by using any of the ABI nucleic acid synthesizer series (Applied Biosystems) or other automated systems known in the art. Antisense sequences can also be produced biologically, such as by transforming an appropriate host cell with an expression vector containing the sequence of interest. (See, e.g., Agrawal, supra.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense  
20 sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E., et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J., et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral  
25 vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998)  
30 *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

### Expression

In order to express a biologically active DITHP, the nucleotide sequences encoding DITHP or  
35 fragments thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in

a suitable host. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding DITHP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, supra, Chapters 4, 8, 5 16, and 17; and Ausubel, supra, Chapters 9, 10, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding DITHP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); 10 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal (mammalian) cell systems. (See, e.g., Sambrook, supra; Ausubel, 1995, supra, Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184; Engelhard, E.K. et al. (1994) 15 *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, 20 J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al., (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. 25 (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.) The invention is not limited by the host cell employed.

For long term production of recombinant proteins in mammalian systems, stable expression of DITHP in cell lines is preferred. For example, sequences encoding DITHP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous 30 expression elements and a selectable marker gene on the same or on a separate vector. Any number of selection systems may be used to recover transformed cell lines. (See, e.g., Wigler, M. et al. (1977) *Cell* 11:223-232; Lowy, I. et al. (1980) *Cell* 22:817-823.; Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbere-Garapin, F. et al. (1981) *J. Mol. Biol.* 150:1-14; Hartman, S.C. and R.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051; Rhodes, C.A. (1995) 35 *Methods Mol. Biol.* 55:121-131.)

Therapeutic Uses of dithp

The dithp of the invention may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and Somia, N. (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in dithp expression or regulation causes disease, the expression of dithp from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in dithp are treated by constructing mammalian expression vectors comprising dithp and introducing these vectors by mechanical means into dithp-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and Anderson, W.F. (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and Récipon, H. (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of dithp include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). The dithp of the invention may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and Bujard, H. (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen, M. et al., (1995) Science 268:1766-1769; Rossi, F.M.V. and Blau, H.M. (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX

plasmid (Invitrogen); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding DITHP from a normal individual.

- 5           Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and Eb, A.J. (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. 10 (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

- In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to dithp expression are treated by constructing a retrovirus vector consisting of (i) dithp under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) 15 appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that 20 expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and Miller, A.D. (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high 25 transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; 30 Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

- In the alternative, an adenovirus-based gene therapy delivery system is used to deliver dithp to cells which have one or more genetic abnormalities with respect to the expression of dithp. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill 35 in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995)

Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and Somia, N. (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

5 In another alternative, a herpes-based, gene therapy delivery system is used to deliver dithp to target cells which have one or more genetic abnormalities with respect to the expression of dithp. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing dithp to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a  
10 reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists  
15 of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W. F. et al. 1999 *J. Virol.* 73:519-532 and Xu, H. et al., (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences,  
20 the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver dithp to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV),  
25 has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and Li, K-J. (1998) *Curr. Opin. Biotech.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase).  
30 Similarly, inserting dithp into the alphavirus genome in place of the capsid-coding region results in the production of a large number of dithp RNAs and the synthesis of high levels of DITHP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the  
35 needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of dithp into a variety of cell types. The specific

transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

5

#### Antibodies

Anti-DITHP antibodies may be used to analyze protein expression levels. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, and Fab fragments. For descriptions of and protocols of antibody technologies, see, e.g., Pound J.D. (1998)

10 Immunochemical Protocols, Humana Press, Totowa, NJ.

The amino acid sequence encoded by the dithp of the Sequence Listing may be analyzed by appropriate software (e.g., LASERGENE NAVIGATOR software, DNASTAR) to determine regions of high immunogenicity. The optimal sequences for immunization are selected from the C-terminus, the N-terminus, and those intervening, hydrophilic regions of the polypeptide which are likely to be  
15 exposed to the external environment when the polypeptide is in its natural conformation. Analysis used to select appropriate epitopes is also described by Ausubel (1997, supra, Chapter 11.7). Peptides used for antibody induction do not need to have biological activity; however, they must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids, and most preferably at least 15 amino  
20 acids. A peptide which mimics an antigenic fragment of the natural polypeptide may be fused with another protein such as keyhole limpet hemocyanin (KLH; Sigma, St. Louis MO) for antibody production. A peptide encompassing an antigenic region may be expressed from a dithp, synthesized as described above, or purified from human cells.

Procedures well known in the art may be used for the production of antibodies. Various hosts  
25 including mice, goats, and rabbits, may be immunized by injection with a peptide. Depending on the host species, various adjuvants may be used to increase immunological response.

In one procedure, peptides about 15 residues in length may be synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using fmoc-chemistry and coupled to KLH (Sigma) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (Ausubel, 1995, supra). Rabbits  
30 are immunized with the peptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for anti-peptide activity by binding the peptide to plastic, blocking with 1% bovine serum albumin (BSA), reacting with rabbit antisera, washing, and reacting with radioiodinated goat anti-rabbit IgG. Antisera with anti-peptide activity are tested for anti-DITHP activity using protocols well known in the art, including ELISA, radioimmunoassay (RIA), and immunoblotting.

35 In another procedure, isolated and purified peptide may be used to immunize mice (about 100  $\mu$ g of peptide) or rabbits (about 1 mg of peptide). Subsequently, the peptide is radioiodinated and

used to screen the immunized animals' B-lymphocytes for production of antipeptide antibodies. Positive cells are then used to produce hybridomas using standard techniques. About 20 mg of peptide is sufficient for labeling and screening several thousand clones. Hybridomas of interest are detected by screening with radioiodinated peptide to identify those fusions producing peptide-specific  
5 monoclonal antibody. In a typical protocol, wells of a multi-well plate (FAST, Becton-Dickinson, Palo Alto, CA) are coated with affinity-purified, specific rabbit-anti-mouse (or suitable anti-species IgG) antibodies at 10 mg/ml. The coated wells are blocked with 1% BSA and washed and exposed to supernatants from hybridomas. After incubation, the wells are exposed to radiolabeled peptide at 1 mg/ml.

10 Clones producing antibodies bind a quantity of labeled peptide that is detectable above background. Such clones are expanded and subjected to 2 cycles of cloning. Cloned hybridomas are injected into pristane-treated mice to produce ascites, and monoclonal antibody is purified from the ascitic fluid by affinity chromatography on protein A (Amersham Pharmacia Biotech). Several procedures for the production of monoclonal antibodies, including *in vitro* production, are described  
15 in Pound (*supra*). Monoclonal antibodies with antipeptide activity are tested for anti-DITHP activity using protocols well known in the art, including ELISA, RIA, and immunoblotting.

Antibody fragments containing specific binding sites for an epitope may also be generated. For example, such fragments include, but are not limited to, the F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule, and the Fab fragments generated by reducing the disulfide bridges  
20 of the F(ab')<sub>2</sub> fragments. Alternatively, construction of Fab expression libraries in filamentous bacteriophage allows rapid and easy identification of monoclonal fragments with desired specificity (Pound, *supra*, Chaps. 45-47). Antibodies generated against polypeptide encoded by dithp can be used to purify and characterize full-length DITHP protein and its activity, binding partners, etc.

#### 25 Assays Using Antibodies

Anti-DITHP antibodies may be used in assays to quantify the amount of DITHP found in a particular human cell. Such assays include methods utilizing the antibody and a label to detect expression level under normal or disease conditions. The peptides and antibodies of the invention may be used with or without modification or labeled by joining them, either covalently or  
30 noncovalently, with a reporter molecule.

Protocols for detecting and measuring protein expression using either polyclonal or monoclonal antibodies are well known in the art. Examples include ELISA, RIA, and fluorescent activated cell sorting (FACS). Such immunoassays typically involve the formation of complexes between the DITHP and its specific antibody and the measurement of such complexes. These and  
35 other assays are described in Pound (*supra*).

Without further elaboration, it is believed that one skilled in the art can, using the preceding



description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/230,517, U.S. Ser. No. 60/230,599, U.S. Ser. No. 60/230,514, U.S. Ser. No. 60/231,167, U.S. Ser. No. 60/230,598, U.S. Ser. No. 60/230,988, U.S. Ser. No. 60/230,518, U.S. Ser. No. 60/230,515, U.S. Ser. No. 60/229,751, U.S. Ser. No. 60/230,610, U.S. Ser. No. 60/229,749, U.S. Ser. No. 60/229,750, U.S. Ser. No. 60/230,597, U.S. Ser. No. 60/230,505, U.S. Ser. No. 60/231,163, U.S. Ser. No. 60/229,747, U.S. Ser. No. 60/229,748, U.S. Ser. No. 60/230,583, U.S. Ser. No. 60/230,519, U.S. Ser. No. 60/230,595, U.S. Ser. No. 60/230,865, and U.S. Ser. No. 60/230,951, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from CLONTECH Laboratories, Inc. (Palo Alto CA) or isolated from various tissues. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In most cases, RNA was treated with DNase. For most libraries, poly(A<sup>+</sup>) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega Corporation (Promega), Madison WI), OLIGOTEX latex particles (QIAGEN, Inc. (QIAGEN), Valencia CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Inc., Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene Cloning Systems, Inc. (Stratagene), La Jolla CA) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, Chapters 5.1 through 6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction

enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof.

Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: the Magic or WIZARD Minipreps DNA purification system (Promega); the AGTC Miniprep purification kit (Edge BioSystems, Gaithersburg MD); and the QIAWELL 8, QIAWELL 8 Plus, and QIAWELL 8 Ultra plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit (QIAGEN). Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format. (Rao, V.B. (1994) Anal. Biochem. 216:1-14.) Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Inc. (Molecular Probes), Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 thermal cycler (Applied Biosystems) or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific Corp., Sunnyvale CA) or the MICROLAB 2200 liquid transfer system (Hamilton). cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, Chapter 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

#### IV. Assembly and Analysis of Sequences

Component sequences from chromatograms were subject to PHRED analysis and assigned a quality score. The sequences having at least a required quality score were subject to various pre-processing editing pathways to eliminate, e.g., low quality 3' ends, vector and linker sequences, polyA  
5 tails, Alu repeats, mitochondrial and ribosomal sequences, bacterial contamination sequences, and sequences smaller than 50 base pairs. In particular, low-information sequences and repetitive elements (e.g., dinucleotide repeats, Alu repeats, etc.) were replaced by "n's", or masked, to prevent spurious matches.

Processed sequences were then subject to assembly procedures in which the sequences were  
10 assigned to gene bins (bins). Each sequence could only belong to one bin. Sequences in each gene bin were assembled to produce consensus sequences (templates). Subsequent new sequences were added to existing bins using BLASTn (v.1.4 WashU) and CROSSMATCH. Candidate pairs were identified as all BLAST hits having a quality score greater than or equal to 150. Alignments of at least 82% local identity were accepted into the bin. The component sequences from each bin were  
15 assembled using a version of PHRAP. Bins with several overlapping component sequences were assembled using DEEP PHRAP. The orientation (sense or antisense) of each assembled template was determined based on the number and orientation of its component sequences. Template sequences as disclosed in the sequence listing correspond to sense strand sequences (the "forward" reading frames), to the best determination. The complementary (antisense) strands are inherently disclosed  
20 herein. The component sequences which were used to assemble each template consensus sequence are listed in Table 5, along with their positions along the template nucleotide sequences.

Bins were compared against each other and those having local similarity of at least 82% were combined and reassembled. Reassembled bins having templates of insufficient overlap (less than 95% local identity) were re-split. Assembled templates were also subject to analysis by  
25 STITCHER/EXON MAPPER algorithms which analyze the probabilities of the presence of splice variants, alternatively spliced exons, splice junctions, differential expression of alternative spliced genes across tissue types or disease states, etc. These resulting bins were subject to several rounds of the above assembly procedures.

Once gene bins were generated based upon sequence alignments, bins were clone joined  
30 based upon clone information. If the 5' sequence of one clone was present in one bin and the 3' sequence from the same clone was present in a different bin, it was likely that the two bins actually belonged together in a single bin. The resulting combined bins underwent assembly procedures to regenerate the consensus sequences.

The final assembled templates were subsequently annotated using the following procedure.  
35 Template sequences were analyzed using BLASTn (v2.0, NCBI) versus gbpri (GenBank version 124). "Hits" were defined as an exact match having from 95% local identity over 200 base pairs

through 100% local identity over 100 base pairs, or a homolog match having an E-value, i.e. a probability score, of  $\leq 1 \times 10^{-8}$ . The hits were subject to frameshift FASTx versus GENPEPT (GenBank version 124). (See Table 8). In this analysis, a homolog match was defined as having an E-value of  $\leq 1 \times 10^{-8}$ . The assembly method used above was described in "System and Methods for  
5 Analyzing Biomolecular Sequences," U.S.S.N. 09/276,534, filed March 25, 1999, and the LIFESEQ Gold user manual (Incyte) both incorporated by reference herein.

Following assembly, template sequences were subjected to motif, BLAST, and functional analyses, and categorized in protein hierarchies using methods described in, e.g., "Database System Employing Protein Function Hierarchies for Viewing Biomolecular Sequence Data," U.S.S.N.  
10 08/812,290, filed March 6, 1997; "Relational Database for Storing Biomolecule Information," U.S.S.N. 08/947,845, filed October 9, 1997; "Project-Based Full-Length Biomolecular Sequence Database," U.S.S.N. 08/811,758, filed March 6, 1997; and "Relational Database and System for Storing Information Relating to Biomolecular Sequences," U.S.S.N. 09/034,807, filed March 4, 1998, all of which are incorporated by reference herein.

15 The template sequences were further analyzed by translating each template in all three forward reading frames and searching each translation against the Pfam database of hidden Markov model-based protein families and domains using the HMMER software package (available to the public from Washington University School of Medicine, St. Louis MO). Regions of templates which, when translated, contain similarity to Pfam consensus sequences are reported in Table 3, along with  
20 descriptions of Pfam protein domains and families. Only those Pfam hits with an E-value of  $\leq 1 \times 10^{-3}$  are reported. (See also World Wide Web site <http://pfam.wustl.edu/> for detailed descriptions of Pfam protein domains and families.)

Additionally, the template sequences were translated in all three forward reading frames, and each translation was searched against hidden Markov models for signal peptides using the HMMER  
25 software package. Construction of hidden Markov models and their usage in sequence analysis has been described. (See, for example, Eddy, S.R. (1996) Curr. Opin. Str. Biol. 6:361-365.) Only those signal peptide hits with a cutoff score of 11 bits or greater are reported. A cutoff score of 11 bits or greater corresponds to at least about 91-94% true-positives in signal peptide prediction. Template sequences were also translated in all three forward reading frames, and each translation was searched  
30 against TMAP, a program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation, with respect to the cell cytosol (Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371). Regions of templates which, when translated, contain similarity to signal peptide or transmembrane consensus sequences are reported in Table 4.

35 The results of HMMER analysis as reported in Tables 3 and 4 may support the results of BLAST analysis as reported in Table 2 or may suggest alternative or additional properties of

template-encoded polypeptides not previously uncovered by BLAST or other analyses.

Template sequences are further analyzed using the bioinformatics tools listed in Table 8, or using sequence analysis software known in the art such as MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR).

- 5 Template sequences may be further queried against public databases such as the GenBank rodent, mammalian, vertebrate, prokaryote, and eukaryote databases.

The template sequences were translated to derive the corresponding longest open reading frame as presented by the polypeptide sequences as reported in Table 7. Alternatively, a polypeptide of the invention may begin at any of the methionine residues within the full length translated  
10 polypeptide. Polypeptide sequences were subsequently analyzed by querying against the GenBank protein database (GENPEPT, (GenBank version 124)). Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the  
15 MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (GENPEPT) database. Column 1 shows the polypeptide sequence identification number (SEQ ID NO:) for the polypeptide segments of the  
20 invention. Column 2 shows the reading frame used in the translation of the polynucleotide sequences encoding the polypeptide segments. Column 3 shows the length of the translated polypeptide segments. Columns 4 and 5 show the start and stop nucleotide positions of the polynucleotide sequences encoding the polypeptide segments. Column 6 shows the GenBank identification number (GI Number) of the nearest GenBank homolog. Column 7 shows the probability score for the match  
25 between each polypeptide and its GenBank homolog. Column 8 shows the annotation of the GenBank homolog.

#### V. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs  
30 from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the  
35 computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

5 The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair  
 10 (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced  
 15 either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

## VI. Tissue Distribution Profiling

20 A tissue distribution profile is determined for each template by compiling the cDNA library tissue classifications of its component cDNA sequences. Each component sequence, is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic  
 25 and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. Template sequences, component sequences, and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

Table 6 shows the tissue distribution profile for the templates of the invention. For each  
 30 template, the three most frequently observed tissue categories are shown in column 3, along with the percentage of component sequences belonging to each category. Only tissue categories with percentage values of  $\geq 10\%$  are shown. A tissue distribution of "widely distributed" in column 3 indicates percentage values of  $< 10\%$  in all tissue categories.

## 35 VII. Transcript Image Analysis

Transcript images are generated as described in Seilhamer et al., "Comparative Gene

Transcript Analysis," U.S. Patent Number 5,840,484, incorporated herein by reference.

### VIII. Extension of Polynucleotide Sequences and Isolation of a Full-length cDNA

Oligonucleotide primers designed using a dithp of the Sequence Listing are used to extend  
5 the nucleic acid sequence. One primer is synthesized to initiate 5' extension of the template, and the  
other primer, to initiate 3' extension of the template. The initial primers may be designed using  
OLIGO 4.06 software (National Biosciences, Inc. (National Biosciences), Plymouth MN), or another  
appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50%  
or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any  
10 stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations are  
avoided. Selected human cDNA libraries are used to extend the sequence. If more than one  
extension is necessary or desired, additional or nested sets of primers are designed.

High fidelity amplification is obtained by PCR using methods well known in the art. PCR is  
performed in 96-well plates using the PTC-200 thermal cycler (MJ Research). The reaction mix  
15 contains DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -  
mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life  
Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair  
PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2  
min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the  
20 alternative, the parameters for primer pair T7 and SK+ are as follows: Step 1: 94°C, 3 min; Step 2:  
94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times;  
Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well is determined by dispensing 100  $\mu$ l PICOGREEN  
quantitation reagent (0.25% (v/v); Molecular Probes) dissolved in 1X Tris-EDTA (TE) and 0.5  $\mu$ l of  
25 undiluted PCR product into each well of an opaque fluorimeter plate (Corning Incorporated  
(Corning), Corning NY), allowing the DNA to bind to the reagent. The plate is scanned in a  
FLUOROSKAN II (Labsystems Oy) to measure the fluorescence of the sample and to quantify the  
concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture is analyzed by electrophoresis  
on a 1 % agarose mini-gel to determine which reactions are successful in extending the sequence.

30 The extended nucleotides are desalted and concentrated, transferred to 384-well plates,  
digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and  
sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For  
shotgun sequencing, the digested nucleotides are separated on low concentration (0.6 to 0.8%)  
agarose gels, fragments are excised, and agar digested with AGAR ACE (Promega). Extended clones  
35 are religated using T4 ligase (New England Biolabs, Inc., Beverly MA) into pUC 18 vector  
(Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction

site overhangs, and transfected into competent *E. coli* cells. Transformed cells are selected on antibiotic-containing media, individual colonies are picked and cultured overnight at 37°C in 384-well plates in LB/2x carbenicillin liquid media.

The cells are lysed, and DNA is amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA is quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries are reamplified using the same conditions as described above. Samples are diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, the dithp is used to obtain regulatory sequences (promoters, introns, and enhancers) using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

#### **IX. Labeling of Probes and Southern Hybridization Analyses**

Hybridization probes derived from the dithp of the Sequence Listing are employed for screening cDNAs, mRNAs, or genomic DNA. The labeling of probe nucleotides between 100 and 1000 nucleotides in length is specifically described, but essentially the same procedure may be used with larger cDNA fragments. Probe sequences are labeled at room temperature for 30 minutes using a T4 polynucleotide kinase,  $\gamma^{32}\text{P}$ -ATP, and 0.5X One-Phor-All Plus (Amersham Pharmacia Biotech) buffer and purified using a ProbeQuant G-50 Microcolumn (Amersham Pharmacia Biotech). The probe mixture is diluted to  $10^7$  dpm/ $\mu\text{g/ml}$  hybridization buffer and used in a typical membrane-based hybridization analysis.

The DNA is digested with a restriction endonuclease such as Eco RV and is electrophoresed through a 0.7% agarose gel. The DNA fragments are transferred from the agarose to nylon membrane (NYTRAN Plus, Schleicher & Schuell, Inc., Keene NH) using procedures specified by the manufacturer of the membrane. Prehybridization is carried out for three or more hours at 68°C, and hybridization is carried out overnight at 68°C. To remove non-specific signals, blots are sequentially washed at room temperature under increasingly stringent conditions, up to 0.1x saline sodium citrate (SSC) and 0.5% sodium dodecyl sulfate. After the blots are placed in a PHOSPHORIMAGER cassette (Molecular Dynamics) or are exposed to autoradiography film, hybridization patterns of standard and experimental lanes are compared. Essentially the same procedure is employed when screening RNA.



## X. Chromosome Mapping of dithp

The cDNA sequences which were used to assemble SEQ ID NO:1-275 are compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that match SEQ ID NO:1-275 are assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as PHRAP (Table 8). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon are used to determine if any of the clustered sequences have been previously mapped. Inclusion of a mapped sequence in a cluster will result in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location. The genetic map locations of SEQ ID NO:1-275 are described as ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters.

## XI. Microarray Analysis

### 20 Probe Preparation from Tissue or Cell Samples

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and polyA<sup>+</sup> RNA is purified using the oligo (dT) cellulose method. Each polyA<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ $\mu$ l oligo-dT primer (21mer), 1X first strand buffer, 0.03 units/ $\mu$ l RNase inhibitor, 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, 40  $\mu$ M dCTP, 40  $\mu$ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng polyA<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control polyA<sup>+</sup> RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA (W. Lei, unpublished). As quantitative controls, the control mRNAs at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng are diluted into reverse transcription reaction at ratios of 1:100,000, 1:10,000, 1:1000, 1:100 (w/w) to sample mRNA respectively. The control mRNAs are diluted into reverse transcription reaction at ratios of 1:3, 3:1, 1:10, 10:1, 1:25, 25:1 (w/w) to sample mRNA differential expression patterns. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Probes are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated

using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The probe is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

#### 5 Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester, PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford, MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

#### Hybridization

Hybridization reactions contain 9  $\mu$ l of probe mixture consisting of 0.2  $\mu$ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The probe mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5x SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

### Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is  
5 focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially.  
10 Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source,  
15 although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the probe mix at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two probes from  
20 different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital  
25 (A/D) conversion board (Analog Devices, Inc., Norwood, MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping  
30 emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

35

## **XII. Complementary Nucleic Acids**

Sequences complementary to the dithp are used to detect, decrease, or inhibit expression of the naturally occurring nucleotide. The use of oligonucleotides comprising from about 15 to 30 base pairs is typical in the art. However, smaller or larger sequence fragments can also be used.

Appropriate oligonucleotides are designed from the dithp using OLIGO 4.06 software (National Biosciences) or other appropriate programs and are synthesized using methods standard in the art or ordered from a commercial supplier. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent transcription factor binding to the promoter sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding and processing of the transcript.

### XIII. Expression of DITHP

Expression and purification of DITHP is accomplished using bacterial or virus-based expression systems. For expression of DITHP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express DITHP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of DITHP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding DITHP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See e.g., Engelhard, supra; and Sandig, supra.)

In most expression systems, DITHP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from DITHP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak Company, Rochester NY). 6-His, a stretch of six consecutive histidine residues, enables

purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, Chapters 10 and 16). Purified DITHP obtained by these methods can be used directly in the following activity assay.

#### 5 XIV. Demonstration of DITHP Activity

DITHP activity is demonstrated through a variety of specific assays, some of which are outlined below.

Oxidoreductase activity of DITHP is measured by the increase in extinction coefficient of NAD(P)H coenzyme at 340 nm for the measurement of oxidation activity, or the decrease in  
10 extinction coefficient of NAD(P)H coenzyme at 340 nm for the measurement of reduction activity (Dalziel, K. (1963) J. Biol. Chem. 238:2850-2858). One of three substrates may be used: Asn- $\beta$ Gal, biocytidine, or ubiquinone-10. The respective subunits of the enzyme reaction, for example, cytochrome  $c_1$ -b oxidoreductase and cytochrome c, are reconstituted. The reaction mixture contains a) 1-2 mg/ml DITHP; and b) 15 mM substrate, 2.4 mM NAD(P)<sup>+</sup> in 0.1 M phosphate buffer, pH 7.1  
15 (oxidation reaction), or 2.0 mM NAD(P)H, in 0.1 M Na<sub>2</sub>HPO<sub>4</sub> buffer, pH 7.4 (reduction reaction); in a total volume of 0.1 ml. Changes in absorbance at 340 nm ( $A_{340}$ ) are measured at 23.5° C using a recording spectrophotometer (Shimadzu Scientific Instruments, Inc., Pleasanton CA). The amount of NAD(P)H is stoichiometrically equivalent to the amount of substrate initially present, and the change in  $A_{340}$  is a direct measure of the amount of NAD(P)H produced;  $\Delta A_{340} = 6620[\text{NADH}]$ .  
20 Oxidoreductase activity of DITHP activity is proportional to the amount of NAD(P)H present in the assay.

Transferase activity of DITHP is measured through assays such as a methyl transferase assay in which the transfer of radiolabeled methyl groups between a donor substrate and an acceptor substrate is measured (Bokar, J.A. et al. (1994) J. Biol. Chem. 269:17697-17704). Reaction mixtures  
25 (50  $\mu$ l final volume) contain 15 mM HEPES, pH 7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 3% polyvinylalcohol, 1.5  $\mu$ Ci [*methyl*-<sup>3</sup>H]AdoMet (0.375  $\mu$ M AdoMet) (DuPont-NEN), 0.6  $\mu$ g DITHP, and acceptor substrate (0.4  $\mu$ g [<sup>35</sup>S]RNA or 6-mercaptapurine (6-MP) to 1 mM final concentration). Reaction mixtures are incubated at 30 °C for 30 minutes, then 65 °C for 5 minutes. The products are separated by chromatography or electrophoresis and the level of methyl transferase activity is  
30 determined by quantification of *methyl*-<sup>3</sup>H recovery.

DITHP hydrolase activity is measured by the hydrolysis of appropriate synthetic peptide substrates conjugated with various chromogenic molecules in which the degree of hydrolysis is quantified by spectrophotometric (or fluorometric) absorption of the released chromophore. (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New  
35 York NY, pp. 25-55) Peptide substrates are designed according to the category of protease activity as endopeptidase (serine, cysteine, aspartic proteases), aminopeptidase (leucine aminopeptidase), or

carboxypeptidase (Carboxypeptidase A and B, procollagen C-proteinase).

DITHP isomerase activity such as peptidyl prolyl *cis/trans* isomerase activity can be assayed by an enzyme assay described by Rahfeld, J.U., et al. (1994) (FEBS Lett. 352: 180-184). The assay is performed at 10°C in 35 mM HEPES buffer, pH 7.8, containing chymotrypsin (0.5 mg/ml) and

5 DITHP at a variety of concentrations. Under these assay conditions, the substrate, Suc-Ala-Xaa-Pro-Phe-4-NA, is in equilibrium with respect to the prolyl bond, with 80-95% in *trans* and 5-20% in *cis* conformation. An aliquot (2 ul) of the substrate dissolved in dimethyl sulfoxide (10 mg/ml) is added to the reaction mixture described above. Only the *cis* isomer of the substrate is a substrate for cleavage by chymotrypsin. Thus, as the substrate is isomerized by DITHP, the product is cleaved by  
10 chymotrypsin to produce 4-nitroanilide, which is detected by its absorbance at 390 nm. 4-Nitroanilide appears in a time-dependent and a DITHP concentration-dependent manner.

An assay for DITHP activity associated with growth and development measures cell proliferation as the amount of newly initiated DNA synthesis in Swiss mouse 3T3 cells. A plasmid containing polynucleotides encoding DITHP is transfected into quiescent 3T3 cultured cells using  
15 methods well known in the art. The transiently transfected cells are then incubated in the presence of [<sup>3</sup>H]thymidine, a radioactive DNA precursor. Where applicable, varying amounts of DITHP ligand are added to the transfected cells. Incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA.

20 Growth factor activity of DITHP is measured by the stimulation of DNA synthesis in Swiss mouse 3T3 cells (McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY). Initiation of DNA synthesis indicates the cells' entry into the mitotic cycle and their commitment to undergo later division. 3T3 cells are competent to respond to most growth factors, not only those that are mitogenic, but also those that are involved in embryonic  
25 induction. This competence is possible because the *in vivo* specificity demonstrated by some growth factors is not necessarily inherent but is determined by the responding tissue. In this assay, varying amounts of DITHP are added to quiescent 3T3 cultured cells in the presence of [<sup>3</sup>H]thymidine, a radioactive DNA precursor. DITHP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA is measured  
30 over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold DITHP concentration range is indicative of growth factor activity. One unit of activity per milliliter is defined as the concentration of DITHP producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA.

35 Alternatively, an assay for cytokine activity of DITHP measures the proliferation of leukocytes. In this assay, the amount of tritiated thymidine incorporated into newly synthesized

DNA is used to estimate proliferative activity. Varying amounts of DITHP are added to cultured leukocytes, such as granulocytes, monocytes, or lymphocytes, in the presence of [<sup>3</sup>H]thymidine, a radioactive DNA precursor. DITHP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA is measured  
5 over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold DITHP concentration range is indicative of DITHP activity. One unit of activity per milliliter is conventionally defined as the concentration of DITHP producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA.

10 An alternative assay for DITHP cytokine activity utilizes a Boyden micro chamber (Neuroprobe, Cabin John MD) to measure leukocyte chemotaxis (Vicari, supra). In this assay, about 10<sup>5</sup> migratory cells such as macrophages or monocytes are placed in cell culture media in the upper compartment of the chamber. Varying dilutions of DITHP are placed in the lower compartment. The two compartments are separated by a 5 or 8 micron pore polycarbonate filter (Nucleopore, Pleasanton  
15 CA). After incubation at 37 °C for 80 to 120 minutes, the filters are fixed in methanol and stained with appropriate labeling agents. Cells which migrate to the other side of the filter are counted using standard microscopy. The chemotactic index is calculated by dividing the number of migratory cells counted when DITHP is present in the lower compartment by the number of migratory cells counted when only media is present in the lower compartment. The chemotactic index is proportional to the  
20 activity of DITHP.

Alternatively, cell lines or tissues transformed with a vector containing dithp can be assayed for DITHP activity by immunoblotting. Cells are denatured in SDS in the presence of β-mercaptoethanol, nucleic acids removed by ethanol precipitation, and proteins purified by acetone precipitation. Pellets are resuspended in 20 mM tris buffer at pH 7.5 and incubated with Protein G-  
25 Sepharose pre-coated with an antibody specific for DITHP. After washing, the Sepharose beads are boiled in electrophoresis sample buffer, and the eluted proteins subjected to SDS-PAGE. The SDS-PAGE is transferred to a nitrocellulose membrane for immunoblotting, and the DITHP activity is assessed by visualizing and quantifying bands on the blot using the antibody specific for DITHP as the primary antibody and <sup>125</sup>I-labeled IgG specific for the primary antibody as the secondary antibody.

30 DITHP kinase activity is measured by phosphorylation of a protein substrate using γ-labeled [<sup>32</sup>P]-ATP and quantitation of the incorporated radioactivity using a radioisotope counter. DITHP is incubated with the protein substrate, [<sup>32</sup>P]-ATP, and an appropriate kinase buffer. The [<sup>32</sup>P] incorporated into the product is separated from free [<sup>32</sup>P]-ATP by electrophoresis and the incorporated [<sup>32</sup>P] is counted. The amount of [<sup>32</sup>P] recovered is proportional to the kinase activity of  
35 DITHP in the assay. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

In the alternative, DITHP activity is measured by the increase in cell proliferation resulting from transformation of a mammalian cell line such as COS7, HeLa or CHO with an eukaryotic expression vector encoding DITHP. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. The cells are  
5 incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression of DITHP. Phase microscopy is then used to compare the mitotic index of transformed versus control cells. An increase in the mitotic index indicates DITHP activity.

In a further alternative, an assay for DITHP signaling activity is based upon the ability of GPCR family proteins to modulate G protein-activated second messenger signal transduction  
10 pathways (e.g., cAMP; Gaudin, P. et al. (1998) J. Biol. Chem. 273:4990-4996). A plasmid encoding full length DITHP is transfected into a mammalian cell line (e.g., Chinese hamster ovary (CHO) or human embryonic kidney (HEK-293) cell lines) using methods well-known in the art. Transfected cells are grown in 12-well trays in culture medium for 48 hours, then the culture medium is discarded, and the attached cells are gently washed with PBS. The cells are then incubated in culture  
15 medium with or without ligand for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from cells exposed to ligand compared to those without ligand are proportional to the amount of DITHP present in the transfected cells.

Alternatively, an assay for DITHP protein phosphatase activity measures the hydrolysis of P-nitrophenyl phosphate (PNPP). DITHP is incubated together with PNPP in HEPES buffer pH 7.5, in the presence of 0.1%  $\beta$ -mercaptoethanol at 37°C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH, and the increase in light absorbance of the reaction mixture at 410 nm  
20 resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the phosphatase activity of DITHP in the assay (Diamond, R.H. et al (1994) Mol Cell Biol 14:3752-3762).

An alternative assay measures DITHP-mediated G-protein signaling activity by monitoring the mobilization of  $Ca^{++}$  as an indicator of the signal transduction pathway stimulation. (See, e.g., Grynkievich, G. et al. (1985) J. Biol. Chem. 260:3440; McColl, S. et al. (1993) J. Immunol.  
30 150:4550-4555; and Aussel, C. et al. (1988) J. Immunol. 140:215-220). The assay requires preloading neutrophils or T cells with a fluorescent dye such as FURA-2 or BCECF (Universal Imaging Corp, Westchester PA) whose emission characteristics are altered by  $Ca^{++}$  binding. When the cells are exposed to one or more activating stimuli artificially (e.g., anti-CD3 antibody ligation of the T cell receptor) or physiologically (e.g., by allogeneic stimulation),  $Ca^{++}$  flux takes place. This  
35 flux can be observed and quantified by assaying the cells in a fluorometer or fluorescent activated cell sorter. Measurements of  $Ca^{++}$  flux are compared between cells in their normal state and those



transfected with DITHP. Increased  $\text{Ca}^{++}$  mobilization attributable to increased DITHP concentration is proportional to DITHP activity.

DITHP transport activity is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with DITHP mRNA (10 ng per oocyte) and  
5 incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM  $\text{CaCl}_2$ , 1mM  $\text{MgCl}_2$ , 1mM  $\text{Na}_2\text{HPO}_4$ , 5 mM Hepes, 3.8 mM NaOH, 50µg/ml gentamycin, pH 7.8) to allow expression of DITHP protein. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM  $\text{CaCl}_2$ , 1mM  $\text{MgCl}_2$ , 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled  
10 substrate (e.g. radiolabeled with  $^3\text{H}$ , fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in  $\text{Na}^+$ -free medium, measuring the incorporated label, and comparing with controls. DITHP transport activity is proportional to the level of internalized labeled substrate.

DITHP transferase activity is demonstrated by a test for galactosyltransferase activity. This  
15 can be determined by measuring the transfer of radiolabeled galactose from UDP-galactose to a GlcNAc-terminated oligosaccharide chain (Kolbinger, F. et al. (1998) J. Biol. Chem. 273:58-65). The sample is incubated with 14 µl of assay stock solution (180 mM sodium cacodylate, pH 6.5, 1 mg/ml bovine serum albumin, 0.26 mM UDP-galactose, 2 µl of UDP- $^3\text{H}$ galactose), 1 µl of  $\text{MnCl}_2$  (500 mM), and 2.5 µl of  $\text{GlcNAc}\beta\text{O}-(\text{CH}_2)_6-\text{CO}_2\text{Me}$  (37 mg/ml in dimethyl sulfoxide) for 60 minutes  
20 at 37°C. The reaction is quenched by the addition of 1 ml of water and loaded on a C18 Sep-Pak cartridge (Waters), and the column is washed twice with 5 ml of water to remove unreacted UDP- $^3\text{H}$ galactose. The  $^3\text{H}$ galactosylated  $\text{GlcNAc}\beta\text{O}-(\text{CH}_2)_6-\text{CO}_2\text{Me}$  remains bound to the column during the water washes and is eluted with 5 ml of methanol. Radioactivity in the eluted material is measured by liquid scintillation counting and is proportional to galactosyltransferase activity in the  
25 starting sample.

In the alternative, DITHP induction by heat or toxins may be demonstrated using primary cultures of human fibroblasts or human cell lines such as CCL-13, HEK293, or HEP G2 (ATCC). To heat induce DITHP expression, aliquots of cells are incubated at 42 °C for 15, 30, or 60 minutes. Control aliquots are incubated at 37 °C for the same time periods. To induce DITHP expression by  
30 toxins, aliquots of cells are treated with 100 µM arsenite or 20 mM azetidine-2-carboxylic acid for 0, 3, 6, or 12 hours. After exposure to heat, arsenite, or the amino acid analogue, samples of the treated cells are harvested and cell lysates prepared for analysis by western blot. Cells are lysed in lysis buffer containing 1% Nonidet P-40, 0.15 M NaCl, 50 mM Tris-HCl, 5 mM EDTA, 2 mM N-ethylmaleimide, 2 mM phenylmethylsulfonyl fluoride, 1 mg/ml leupeptin, and 1 mg/ml pepstatin.  
35 Twenty micrograms of the cell lysate is separated on an 8% SDS-PAGE gel and transferred to a membrane. After blocking with 5% nonfat dry milk/phosphate-buffered saline for 1 h, the membrane

is incubated overnight at 4°C or at room temperature for 2-4 hours with a 1:1000 dilution of anti-DITHP serum in 2% nonfat dry milk/phosphate-buffered saline. The membrane is then washed and incubated with a 1:1000 dilution of horseradish peroxidase-conjugated goat anti-rabbit IgG in 2% dry milk/phosphate-buffered saline. After washing with 0.1% Tween 20 in phosphate-buffered saline, the DITHP protein is detected and compared to controls using chemiluminescence.

Alternatively, DITHP protease activity is measured by the hydrolysis of appropriate synthetic peptide substrates conjugated with various chromogenic molecules in which the degree of hydrolysis is quantified by spectrophotometric (or fluorometric) absorption of the released chromophore (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York, NY, pp.25-55). Peptide substrates are designed according to the category of protease activity as endopeptidase (serine, cysteine, aspartic proteases, or metalloproteases), aminopeptidase (leucine aminopeptidase), or carboxypeptidase (carboxypeptidases A and B, procollagen C-proteinase). Commonly used chromogens are 2-naphthylamine, 4-nitroaniline, and furylacrylic acid. Assays are performed at ambient temperature and contain an aliquot of the enzyme and the appropriate substrate in a suitable buffer. Reactions are carried out in an optical cuvette, and the increase/decrease in absorbance of the chromogen released during hydrolysis of the peptide substrate is measured. The change in absorbance is proportional to the DITHP protease activity in the assay.

In the alternative, an assay for DITHP protease activity takes advantage of fluorescence resonance energy transfer (FRET) that occurs when one donor and one acceptor fluorophore with an appropriate spectral overlap are in close proximity. A flexible peptide linker containing a cleavage site specific for PRTS is fused between a red-shifted variant (RSGFP4) and a blue variant (BFP5) of Green Fluorescent Protein. This fusion protein has spectral properties that suggest energy transfer is occurring from BFP5 to RSGFP4. When the fusion protein is incubated with DITHP, the substrate is cleaved, and the two fluorescent proteins dissociate. This is accompanied by a marked decrease in energy transfer which is quantified by comparing the emission spectra before and after the addition of DITHP (Mitra, R.D. et al (1996) *Gene* 173:13-17). This assay can also be performed in living cells. In this case the fluorescent substrate protein is expressed constitutively in cells and DITHP is introduced on an inducible vector so that FRET can be monitored in the presence and absence of DITHP (Sagot, I. et al (1999) *FEBS Lett.* 447:53-57).

A method to determine the nucleic acid binding activity of DITHP involves a polyacrylamide gel mobility-shift assay. In preparation for this assay, DITHP is expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector containing DITHP cDNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of DITHP. Extracts containing solubilized proteins can be prepared from cells expressing DITHP by methods well known in the art.

Portions of the extract containing DITHP are added to [ $^{32}\text{P}$ ]-labeled RNA or DNA. Radioactive nucleic acid can be synthesized in vitro by techniques well known in the art. The mixtures are incubated at 25°C in the presence of RNase- and DNase-inhibitors under buffered conditions for 5-10 minutes. After incubation, the samples are analyzed by polyacrylamide gel electrophoresis followed by autoradiography. The presence of a band on the autoradiogram indicates the formation of a complex between DITHP and the radioactive transcript. A band of similar mobility will not be present in samples prepared using control extracts prepared from untransformed cells.

In the alternative, a method to determine the methylase activity of a DITHP measures transfer of radiolabeled methyl groups between a donor substrate and an acceptor substrate. Reaction mixtures (50  $\mu\text{l}$  final volume) contain 15 mM HEPES, pH 7.9, 1.5 mM  $\text{MgCl}_2$ , 10 mM dithiothreitol, 3% polyvinylalcohol, 1.5  $\mu\text{Ci}$  [*methyl*- $^3\text{H}$ ]AdoMet (0.375  $\mu\text{M}$  AdoMet) (DuPont-NEN), 0.6  $\mu\text{g}$  DITHP, and acceptor substrate (e.g., 0.4  $\mu\text{g}$  [ $^{35}\text{S}$ ]RNA, or 6-mercaptopurine (6-MP) to 1 mM final concentration). Reaction mixtures are incubated at 30°C for 30 minutes, then 65°C for 5 minutes. Analysis of [*methyl*- $^3\text{H}$ ]RNA is as follows: 1) 50  $\mu\text{l}$  of 2 x loading buffer (20 mM Tris-HCl, pH 7.6, 1 M LiCl, 1 mM EDTA, 1% sodium dodecyl sulphate (SDS)) and 50  $\mu\text{l}$  oligo d(T)-cellulose (10 mg/ml in 1 x loading buffer) are added to the reaction mixture, and incubated at ambient temperature with shaking for 30 minutes. 2) Reaction mixtures are transferred to a 96-well filtration plate attached to a vacuum apparatus. 3) Each sample is washed sequentially with three 2.4 ml aliquots of 1 x oligo d(T) loading buffer containing 0.5% SDS, 0.1% SDS, or no SDS. and 4) RNA is eluted with 300  $\mu\text{l}$  of water into a 96-well collection plate, transferred to scintillation vials containing liquid scintillant, and radioactivity determined. Analysis of [*methyl*- $^3\text{H}$ ]6-MP is as follows: 1) 500  $\mu\text{l}$  0.5 M borate buffer, pH 10.0, and then 2.5 ml of 20% (v/v) isoamyl alcohol in toluene are added to the reaction mixtures. 2) The samples mixed by vigorous vortexing for ten seconds. 3) After centrifugation at 700g for 10 minutes, 1.5 ml of the organic phase is transferred to scintillation vials containing 0.5 ml absolute ethanol and liquid scintillant, and radioactivity determined. and 4) Results are corrected for the extraction of 6-MP into the organic phase (approximately 41%).

An assay for adhesion activity of DITHP measures the disruption of cytoskeletal filament networks upon overexpression of DITHP in cultured cell lines (Reznicek, G.A. et al. (1998) J. Cell Biol. 141:209-225). cDNA encoding DITHP is subcloned into a mammalian expression vector that drives high levels of cDNA expression. This construct is transfected into cultured cells, such as rat kangaroo PtK2 or rat bladder carcinoma 804G cells. Actin filaments and intermediate filaments such as keratin and vimentin are visualized by immunofluorescence microscopy using antibodies and techniques well known in the art. The configuration and abundance of cytoskeletal filaments can be assessed and quantified using confocal imaging techniques. In particular, the bundling and collapse of cytoskeletal filament networks is indicative of DITHP adhesion activity.

Alternatively, an assay for DITHP activity measures the expression of DITHP on the cell

surface. cDNA encoding DITHP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using DITHP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of DITHP expressed on the cell surface.

Alternatively, an assay for DITHP activity measures the amount of cell aggregation induced by overexpression of DITHP. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding DITHP contained within a suitable mammalian expression vector under control of a strong promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green Fluorescent Protein (CLONTECH), is useful for identifying stable transfectants. The amount of cell agglutination, or clumping, associated with transfected cells is compared with that associated with untransfected cells. The amount of cell agglutination is a direct measure of DITHP activity.

DITHP may recognize and precipitate antigen from serum. This activity can be measured by the quantitative precipitin reaction (Golub, E.S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland MA, pages 113-115). DITHP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled DITHP. DITHP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable DITHP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable DITHP-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitation curve is obtained, in which the amount of precipitable DITHP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable DITHP-antigen complex is a measure of DITHP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

A microtubule motility assay for DITHP measures motor protein activity. In this assay, recombinant DITHP is immobilized onto a glass slide or similar substrate. Taxol-stabilized bovine brain microtubules (commercially available) in a solution containing ATP and cytosolic extract are perfused onto the slide. Movement of microtubules as driven by DITHP motor activity can be visualized and quantified using video-enhanced light microscopy and image analysis techniques. DITHP motor protein activity is directly proportional to the frequency and velocity of microtubule movement.

Alternatively, an assay for DITHP measures the formation of protein filaments in vitro. A solution of DITHP at a concentration greater than the "critical concentration" for polymer assembly is applied to carbon-coated grids. Appropriate nucleation sites may be supplied in the solution. The grids are negative stained with 0.7% (w/v) aqueous uranyl acetate and examined by electron

microscopy. The appearance of filaments of approximately 25 nm (microtubules), 8 nm (actin), or 10 nm (intermediate filaments) is a demonstration of protein activity.

DITHP electron transfer activity is demonstrated by oxidation or reduction of NADP.

Substrates such as Asn- $\beta$ Gal, biocytidine, or ubiquinone-10 may be used. The reaction mixture  
5 contains 1-2 mg/ml HORP, 15 mM substrate, and 2.4 mM NAD(P)<sup>+</sup> in 0.1 M phosphate buffer, pH 7.1 (oxidation reaction), or 2.0 mM NAD(P)H, in 0.1 M Na<sub>2</sub>HPO<sub>4</sub> buffer, pH 7.4 (reduction reaction); in a total volume of 0.1 ml. FAD may be included with NAD, according to methods well known in the art. Changes in absorbance are measured using a recording spectrophotometer. The amount of NAD(P)H is stoichiometrically equivalent to the amount of substrate initially present, and the change  
10 in A<sub>340</sub> is a direct measure of the amount of NAD(P)H produced;  $\Delta A_{340} = 6620[\text{NADH}]$ . DITHP activity is proportional to the amount of NAD(P)H present in the assay. The increase in extinction coefficient of NAD(P)H coenzyme at 340 nm is a measure of oxidation activity, or the decrease in extinction coefficient of NAD(P)H coenzyme at 340 nm is a measure of reduction activity (Dalziel, K. (1963) J. Biol. Chem. 238:2850-2858).

15 DITHP transcription factor activity is measured by its ability to stimulate transcription of a reporter gene (Liu, H.Y. et al. (1997) EMBO J. 16:5289-5298). The assay entails the use of a well characterized reporter gene construct, LexA<sub>op</sub>-LacZ, that consists of LexA DNA transcriptional control elements (LexA<sub>op</sub>) fused to sequences encoding the *E. coli* LacZ enzyme. The methods for constructing and expressing fusion genes, introducing them into cells, and measuring LacZ enzyme  
20 activity, are well known to those skilled in the art. Sequences encoding DITHP are cloned into a plasmid that directs the synthesis of a fusion protein, LexA-DITHP, consisting of DITHP and a DNA binding domain derived from the LexA transcription factor. The resulting plasmid, encoding a LexA-DITHP fusion protein, is introduced into yeast cells along with a plasmid containing the LexA<sub>op</sub>-LacZ reporter gene. The amount of LacZ enzyme activity associated with LexA-DITHP transfected cells,  
25 relative to control cells, is proportional to the amount of transcription stimulated by the DITHP.

Chromatin activity of DITHP is demonstrated by measuring sensitivity to DNase I (Dawson, B.A. et al. (1989) J. Biol. Chem. 264:12830-12837). Samples are treated with DNase I, followed by insertion of a cleavable biotinylated nucleotide analog, 5-[(N-biotinamido)hexanoamido-ethyl-1,3-thiopropionyl-3-aminoallyl]-2'-deoxyuridine 5'-triphosphate using nick-repair techniques well known  
30 to those skilled in the art. Following purification and digestion with EcoRI restriction endonuclease, biotinylated sequences are affinity isolated by sequential binding to streptavidin and biotincellulose.

Another specific assay demonstrates the ion conductance capacity of DITHP using an electrophysiological assay. DITHP is expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding DITHP. Eukaryotic expression  
35 vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A small amount of a second plasmid, which expresses any one of a number of

marker genes such as  $\beta$ -galactosidase, is co-transformed into the cells in order to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of DITHP and  $\beta$ -galactosidase. Transformed cells expressing  $\beta$ -galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells are tested for differences in membrane conductance due to various ions by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or  $\beta$ -galactosidase sequences alone, are used as controls and tested in parallel. The contribution of DITHP to cation or anion conductance can be shown by incubating the cells using antibodies specific for either DITHP. The respective antibodies will bind to the extracellular side of DITHP, thereby blocking the pore in the ion channel, and the associated conductance.

#### XV. Functional Assays

DITHP function is assessed by expressing dithp at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen Corporation, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected.

Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; CLONTECH), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties.

FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow

Cytometry, Oxford, New York NY.

The influence of DITHP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding DITHP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Inc., Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding DITHP and other genes of interest can be analyzed by northern analysis or microarray techniques.

#### **XVI. Production of Antibodies**

DITHP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the DITHP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding peptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, Chapter 11.)

Typically, peptides 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using fmoc-chemistry and coupled to KLH (Sigma) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, *supra*.) Rabbits are immunized with the peptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated goat anti-rabbit IgG. Antisera with antipeptide activity are tested for anti-DITHP activity using protocols well known in the art, including ELISA, RIA, and immunoblotting.

#### **XVII. Purification of Naturally Occurring DITHP Using Specific Antibodies**

Naturally occurring or recombinant DITHP is substantially purified by immunoaffinity chromatography using antibodies specific for DITHP. An immunoaffinity column is constructed by covalently coupling anti-DITHP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing DITHP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of DITHP (e.g., high ionic strength

buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/DITHP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and DITHP is collected.

5 **XVIII. Identification of Molecules Which Interact with DITHP**

DITHP, or biologically active fragments thereof, are labeled with  $^{125}\text{I}$  Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled DITHP, washed, and any wells with labeled DITHP complex are assayed. Data obtained using different  
10 concentrations of DITHP are used to calculate values for the number, affinity, and association of DITHP with the candidate molecules.

Alternatively, molecules interacting with DITHP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (CLONTECH).

15 DITHP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

20 All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to  
25 such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.



TABLE 1

SEQ ID NO:	Template ID	SEQ ID NO:	ORF ID
1	LG:405741.3:2000SEP08	276	LG:405741.3.orf3:2000SEP08
2	LG:337194.1:2000SEP08	277	LG:337194.1.orf1:2000SEP08
3	LG:017108.4:2000SEP08	278	LG:017108.4.orf2:2000SEP08
4	LG:372569.5:2000SEP08	279	LG:372569.5.orf1:2000SEP08
5	LG:968765.1:2000SEP08	280	LG:968765.1.orf2:2000SEP08
6	LG:255999.16:2000SEP08	281	LG:255999.16.orf3:2000SEP08
7	LG:977820.9:2000SEP08	282	LG:977820.9.orf2:2000SEP08
8	LI:1071608.1:2000SEP08	283	LI:1071608.1.orf3:2000SEP08
9	LI:1074023.1:2000SEP08	284	LI:1074023.1.orf3:2000SEP08
10	LI:453570.1:2000SEP08	285	LI:453570.1.orf3:2000SEP08
11	LI:072072.1:2000SEP08	286	LI:072072.1.orf2:2000SEP08
12	LI:148565.4:2000SEP08	287	LI:148565.4.orf1:2000SEP08
13	LI:368626.4:2000SEP08	288	LI:368626.4.orf1:2000SEP08
14	LI:346123.1:2000SEP08	289	LI:346123.1.orf2:2000SEP08
15	LI:335795.11:2000SEP08	290	LI:335795.11.orf2:2000SEP08
16	LI:246023.2:2000SEP08	291	LI:246023.2.orf1:2000SEP08
17	LG:1100661.1:2000SEP08	292	LG:1100661.1.orf3:2000SEP08
18	LG:475856.1:2000SEP08	293	LG:475856.1.orf1:2000SEP08
19	LG:1015343.1:2000SEP08	294	LG:1015343.1.orf1:2000SEP08
20	LG:1400575.1:2000SEP08	295	LG:1400575.1.orf2:2000SEP08
21	LG:1080545.1:2000SEP08	296	LG:1080545.1.orf2:2000SEP08
22	LG:213947.1:2000SEP08	297	LG:213947.1.orf1:2000SEP08
23	LI:720641.1:2000SEP08	298	LI:720641.1.orf2:2000SEP08
24	LI:1023894.1:2000SEP08	299	LI:1023894.1.orf3:2000SEP08
25	LI:734904.1:2000SEP08	300	LI:734904.1.orf3:2000SEP08
26	LI:1178118.1:2000SEP08	301	LI:1178118.1.orf3:2000SEP08
27	LI:213947.1:2000SEP08	302	LI:213947.1.orf3:2000SEP08
28	LG:407304.1:2000SEP08	303	LG:407304.1.orf1:2000SEP08
29	LG:337358.1:2000SEP08	304	LG:337358.1.orf2:2000SEP08
30	LG:986090.1:2000SEP08	305	LG:986090.1.orf1:2000SEP08
31	LG:123250.1:2000SEP08	306	LG:123250.1.orf2:2000SEP08
32	LG:1028774.2:2000SEP08	307	LG:1028774.2.orf1:2000SEP08
33	LG:338927.6:2000SEP08	308	LG:338927.6.orf2:2000SEP08
34	LG:332944.2:2000SEP08	309	LG:332944.2.orf1:2000SEP08
35	LI:347174.5:2000SEP08	310	LI:347174.5.orf2:2000SEP08
36	LI:477070.1:2000SEP08	311	LI:477070.1.orf2:2000SEP08
37	LI:723144.1:2000SEP08	312	LI:723144.1.orf2:2000SEP08
38	LI:1007188.1:2000SEP08	313	LI:1007188.1.orf1:2000SEP08
39	LI:1024412.1:2000SEP08	314	LI:1024412.1.orf3:2000SEP08
40	LI:284797.3:2000SEP08	315	LI:284797.3.orf3:2000SEP08
41	LI:1092901.1:2000SEP08	316	LI:1092901.1.orf3:2000SEP08
42	LI:228930.1:2000SEP08	317	LI:228930.1.orf2:2000SEP08
43	LI:722913.1:2000SEP08	318	LI:722913.1.orf1:2000SEP08
44	LG:457478.1:2000SEP08	319	LG:457478.1.orf3:2000SEP08
45	LG:358719.1:2000SEP08	320	LG:358719.1.orf3:2000SEP08
46	LG:105160.5:2000SEP08	321	LG:105160.5.orf1:2000SEP08
47	LG:400705.1:2000SEP08	322	LG:400705.1.orf1:2000SEP08
48	LG:221977.1:2000SEP08	323	LG:221977.1.orf3:2000SEP08
49	LG:898771.1:2000SEP08	324	LG:898771.1.orf1:2000SEP08

TABLE 1

SEQ ID NO:	Template ID	SEQ ID NO:	ORF ID
50	LI:457478.1:2000SEP08	325	LI:457478.1.orf2:2000SEP08
51	LI:125140.1:2000SEP08	326	LI:125140.1.orf3:2000SEP08
52	LI:021095.2:2000SEP08	327	LI:021095.2.orf1:2000SEP08
53	LI:888730.1:2000SEP08	328	LI:888730.1.orf2:2000SEP08
53	LI:888730.1:2000SEP08	329	LI:888730.1.orf3:2000SEP08
54	LI:358719.1:2000SEP08	330	LI:358719.1.orf2:2000SEP08
55	LI:351342.3:2000SEP08	331	LI:351342.3.orf3:2000SEP08
56	LI:256099.2:2000SEP08	332	LI:256099.2.orf3:2000SEP08
57	LI:2051991.1:2000SEP08	333	LI:2051991.1.orf1:2000SEP08
58	LG:980769.1:2000SEP08	334	LG:980769.1.orf3:2000SEP08
59	LG:332474.3:2000SEP08	335	LG:332474.3.orf3:2000SEP08
60	LG:1087707.1:2000SEP08	336	LG:1087707.1.orf1:2000SEP08
61	LG:415349.1:2000SEP08	337	LG:415349.1.orf2:2000SEP08
62	LG:132420.2:2000SEP08	338	LG:132420.2.orf2:2000SEP08
63	LG:394201.1:2000SEP08	339	LG:394201.1.orf2:2000SEP08
64	LG:1060884.1:2000SEP08	340	LG:1060884.1.orf3:2000SEP08
65	LG:242191.1:2000SEP08	341	LG:242191.1.orf2:2000SEP08
66	LG:1063762.3:2000SEP08	342	LG:1063762.3.orf1:2000SEP08
67	LG:1100856.1:2000SEP08	343	LG:1100856.1.orf2:2000SEP08
68	LG:979390.2:2000SEP08	344	LG:979390.2.orf1:2000SEP08
69	LG:1400447.1:2000SEP08	345	LG:1400447.1.orf1:2000SEP08
70	LG:1400562.1:2000SEP08	346	LG:1400562.1.orf3:2000SEP08
71	LG:1076130.1:2000SEP08	347	LG:1076130.1.orf3:2000SEP08
72	LG:1064459.1:2000SEP08	348	LG:1064459.1.orf2:2000SEP08
73	LG:1079415.14:2000SEP08	349	LG:1079415.14.orf2:2000SEP08
74	LG:1329431.3:2000SEP08	350	LG:1329431.3.orf2:2000SEP08
75	LG:1088431.2:2000SEP08	351	LG:1088431.2.orf1:2000SEP08
76	LG:1329462.2:2000SEP08	352	LG:1329462.2.orf1:2000SEP08
77	LI:393468.1:2000SEP08	353	LI:393468.1.orf2:2000SEP08
78	LI:722577.1:2000SEP08	354	LI:722577.1.orf1:2000SEP08
79	LI:322783.16:2000SEP08	355	LI:322783.16.orf1:2000SEP08
80	LI:901355.2:2000SEP08	356	LI:901355.2.orf2:2000SEP08
81	LI:038859.2:2000SEP08	357	LI:038859.2.orf1:2000SEP08
82	LI:1046117.1:2000SEP08	358	LI:1046117.1.orf1:2000SEP08
83	LI:801015.1:2000SEP08	359	LI:801015.1.orf1:2000SEP08
84	LI:1175590.1:2000SEP08	360	LI:1175590.1.orf3:2000SEP08
85	LI:1170585.2:2000SEP08	361	LI:1170585.2.orf2:2000SEP08
86	LI:719531.2:2000SEP08	362	LI:719531.2.orf1:2000SEP08
87	LI:794623.1:2000SEP08	363	LI:794623.1.orf2:2000SEP08
88	LI:1173119.1:2000SEP08	364	LI:1173119.1.orf3:2000SEP08
89	LI:1093285.1:2000SEP08	365	LI:1093285.1.orf1:2000SEP08
90	LI:1091881.1:2000SEP08	366	LI:1091881.1.orf1:2000SEP08
91	LI:1091617.1:2000SEP08	367	LI:1091617.1.orf2:2000SEP08
92	LI:1082344.1:2000SEP08	368	LI:1082344.1.orf1:2000SEP08
93	LI:1166249.1:2000SEP08	369	LI:1166249.1.orf3:2000SEP08
94	LI:799675.1:2000SEP08	370	LI:799675.1.orf1:2000SEP08
95	LI:1178899.1:2000SEP08	371	LI:1178899.1.orf2:2000SEP08
96	LI:1169241.1:2000SEP08	372	LI:1169241.1.orf3:2000SEP08
97	LI:1180090.1:2000SEP08	373	LI:1180090.1.orf1:2000SEP08

TABLE 1

SEQ ID NO:	Template ID	SEQ ID NO:	ORF ID
98	LI:2049322.1:2000SEP08	374	LI:2049322.1.orf1:2000SEP08
99	LI:809074.1:2000SEP08	375	LI:809074.1.orf2:2000SEP08
100	LI:805158.1:2000SEP08	376	LI:805158.1.orf2:2000SEP08
101	LI:1172697.1:2000SEP08	377	LI:1172697.1.orf1:2000SEP08
102	LI:1174107.2:2000SEP08	378	LI:1174107.2.orf3:2000SEP08
103	LI:1177434.2:2000SEP08	379	LI:1177434.2.orf2:2000SEP08
104	LI:1184255.1:2000SEP08	380	LI:1184255.1.orf2:2000SEP08
105	LI:1164555.1:2000SEP08	381	LI:1164555.1.orf3:2000SEP08
106	LI:238666.4:2000SEP08	382	LI:238666.4.orf1:2000SEP08
107	LI:1166752.1:2000SEP08	383	LI:1166752.1.orf1:2000SEP08
108	LI:2049654.1:2000SEP08	384	LI:2049654.1.orf3:2000SEP08
109	LI:242665.2:2000SEP08	385	LI:242665.2.orf2:2000SEP08
110	LI:208637.1:2000SEP08	386	LI:208637.1.orf1:2000SEP08
111	LI:2051808.1:2000SEP08	387	LI:2051808.1.orf1:2000SEP08
112	LI:1175136.1:2000SEP08	388	LI:1175136.1.orf2:2000SEP08
113	LI:1177337.1:2000SEP08	389	LI:1177337.1.orf1:2000SEP08
114	LI:1165056.1:2000SEP08	390	LI:1165056.1.orf2:2000SEP08
115	LI:1175250.1:2000SEP08	391	LI:1175250.1.orf1:2000SEP08
116	LI:1183192.1:2000SEP08	392	LI:1183192.1.orf3:2000SEP08
117	LI:1183325.1:2000SEP08	393	LI:1183325.1.orf1:2000SEP08
118	LI:1178269.2:2000SEP08	394	LI:1178269.2.orf3:2000SEP08
119	LI:813422.1:2000SEP08	395	LI:813422.1.orf1:2000SEP08
120	LI:1093049.6:2000SEP08	396	LI:1093049.6.orf2:2000SEP08
121	LI:202192.4:2000SEP08	397	LI:202192.4.orf3:2000SEP08
122	LG:1041854.1:2000SEP08	398	LG:1041854.1.orf3:2000SEP08
123	LG:1100502.1:2000SEP08	399	LG:1100502.1.orf3:2000SEP08
124	LI:726414.1:2000SEP08	400	LI:726414.1.orf1:2000SEP08
125	LI:400517.4:2000SEP08	401	LI:400517.4.orf2:2000SEP08
126	LI:1078917.1:2000SEP08	402	LI:1078917.1.orf2:2000SEP08
127	LI:1012560.1:2000SEP08	403	LI:1012560.1.orf3:2000SEP08
128	LI:427997.4:2000SEP08	404	LI:427997.4.orf3:2000SEP08
129	LI:197899.1:2000SEP08	405	LI:197899.1.orf2:2000SEP08
130	LG:334199.1:2000SEP08	406	LG:334199.1.orf3:2000SEP08
131	LG:334345.1:2000SEP08	407	LG:334345.1.orf2:2000SEP08
132	LG:228092.1:2000SEP08	408	LG:228092.1.orf3:2000SEP08
133	LG:098580.1:2000SEP08	409	LG:098580.1.orf3:2000SEP08
134	LG:969572.1:2000SEP08	410	LG:969572.1.orf3:2000SEP08
135	LG:196958.1:2000SEP08	411	LG:196958.1.orf2:2000SEP08
136	LG:1087811.1:2000SEP08	412	LG:1087811.1.orf2:2000SEP08
137	LG:1327885.1:2000SEP08	413	LG:1327885.1.orf3:2000SEP08
138	LI:449393.1:2000SEP08	414	LI:449393.1.orf3:2000SEP08
139	LI:897616.1:2000SEP08	415	LI:897616.1.orf3:2000SEP08
140	LI:736860.1:2000SEP08	416	LI:736860.1.orf3:2000SEP08
141	LI:027066.6:2000SEP08	417	LI:027066.6.orf3:2000SEP08
142	LI:1074263.1:2000SEP08	418	LI:1074263.1.orf3:2000SEP08
143	LI:334345.1:2000SEP08	419	LI:334345.1.orf1:2000SEP08
144	LI:1093914.1:2000SEP08	420	LI:1093914.1.orf2:2000SEP08
145	LI:1188168.1:2000SEP08	421	LI:1188168.1.orf2:2000SEP08
146	LI:1065168.1:2000SEP08	422	LI:1065168.1.orf1:2000SEP08

TABLE 1

SEQ ID NO:	Template ID	SEQ ID NO:	ORF ID
147	LI:1180418.1:2000SEP08	423	LI:1180418.1.orf2:2000SEP08
148	LG:232648.1:2000SEP08	424	LG:232648.1.orf1:2000SEP08
149	LG:1078420.1:2000SEP08	425	LG:1078420.1.orf2:2000SEP08
150	LG:1397599.1:2000SEP08	426	LG:1397599.1.orf2:2000SEP08
151	LG:1397655.2:2000SEP08	427	LG:1397655.2.orf1:2000SEP08
152	LG:241055.1:2000SEP08	428	LG:241055.1.orf1:2000SEP08
153	LG:1101065.1:2000SEP08	429	LG:1101065.1.orf1:2000SEP08
154	LG:475629.1:2000SEP08	430	LG:475629.1.orf2:2000SEP08
155	LI:348991.1:2000SEP08	431	LI:348991.1.orf1:2000SEP08
156	LI:475629.1:2000SEP08	432	LI:475629.1.orf3:2000SEP08
157	LI:261331.1:2000SEP08	433	LI:261331.1.orf1:2000SEP08
158	LI:815686.1:2000SEP08	434	LI:815686.1.orf3:2000SEP08
159	LI:1167327.2:2000SEP08	435	LI:1167327.2.orf3:2000SEP08
160	LI:758009.3:2000SEP08	436	LI:758009.3.orf2:2000SEP08
161	LG:331593.1:2000SEP08	437	LG:331593.1.orf3:2000SEP08
162	LI:1094174.1:2000SEP08	438	LI:1094174.1.orf2:2000SEP08
163	LI:814362.1:2000SEP08	439	LI:814362.1.orf3:2000SEP08
164	LI:219542.1:2000SEP08	440	LI:219542.1.orf3:2000SEP08
165	LI:726197.1:2000SEP08	441	LI:726197.1.orf3:2000SEP08
166	LI:1075314.1:2000SEP08	442	LI:1075314.1.orf1:2000SEP08
167	LI:437883.1:2000SEP08	443	LI:437883.1.orf2:2000SEP08
168	LG:336265.1:2000SEP08	444	LG:336265.1.orf3:2000SEP08
169	LG:407788.2:2000SEP08	445	LG:407788.2.orf3:2000SEP08
170	LG:1326925.1:2000SEP08	446	LG:1326925.1.orf2:2000SEP08
171	LI:332655.2:2000SEP08	447	LI:332655.2.orf1:2000SEP08
172	LI:1184621.4:2000SEP08	448	LI:1184621.4.orf2:2000SEP08
173	LI:2051386.1:2000SEP08	449	LI:2051386.1.orf2:2000SEP08
174	LG:362757.1:2000SEP08	450	LG:362757.1.orf1:2000SEP08
175	LG:406770.1:2000SEP08	451	LG:406770.1.orf3:2000SEP08
176	LG:1094640.1:2000SEP08	452	LG:1094640.1.orf3:2000SEP08
177	LG:001929.1:2000SEP08	453	LG:001929.1.orf1:2000SEP08
178	LI:401322.1:2000SEP08	454	LI:401322.1.orf3:2000SEP08
179	LI:208748.1:2000SEP08	455	LI:208748.1.orf1:2000SEP08
180	LI:407242.1:2000SEP08	456	LI:407242.1.orf2:2000SEP08
181	LI:403409.1:2000SEP08	457	LI:403409.1.orf3:2000SEP08
182	LI:450798.1:2000SEP08	458	LI:450798.1.orf3:2000SEP08
183	LI:410317.1:2000SEP08	459	LI:410317.1.orf1:2000SEP08
184	LI:340268.1:2000SEP08	460	LI:340268.1.orf3:2000SEP08
185	LI:2051671.1:2000SEP08	461	LI:2051671.1.orf3:2000SEP08
186	LG:998844.1:2000SEP08	462	LG:998844.1.orf3:2000SEP08
187	LG:1043787.1:2000SEP08	463	LG:1043787.1.orf1:2000SEP08
188	LG:1098931.16:2000SEP08	464	LG:1098931.16.orf1:2000SEP08
189	LG:199423.2:2000SEP08	465	LG:199423.2.orf1:2000SEP08
190	LI:1075297.1:2000SEP08	466	LI:1075297.1.orf2:2000SEP08
191	LI:1043321.1:2000SEP08	467	LI:1043321.1.orf3:2000SEP08
192	LI:297070.1:2000SEP08	468	LI:297070.1.orf1:2000SEP08
193	LI:1085041.1:2000SEP08	469	LI:1085041.1.orf3:2000SEP08
194	LI:1071544.1:2000SEP08	470	LI:1071544.1.orf2:2000SEP08
195	LI:2052480.1:2000SEP08	471	LI:2052480.1.orf2:2000SEP08

TABLE 1

SEQ ID NO:	Template ID	SEQ ID NO:	ORF ID
196	LG:450105.1:2000SEP08	472	LG:450105.1.orf2:2000SEP08
197	LG:450581.1:2000SEP08	473	LG:450581.1.orf2:2000SEP08
198	LG:450887.1:2000SEP08	474	LG:450887.1.orf3:2000SEP08
199	LG:460809.1:2000SEP08	475	LG:460809.1.orf3:2000SEP08
200	LG:452089.1:2000SEP08	476	LG:452089.1.orf2:2000SEP08
201	LG:1099416.1:2000SEP08	477	LG:1099416.1.orf3:2000SEP08
202	LG:255713.1:2000SEP08	478	LG:255713.1.orf1:2000SEP08
203	LG:998903.1:2000SEP08	479	LG:998903.1.orf1:2000SEP08
204	LG:1119656.1:2000SEP08	480	LG:1119656.1.orf2:2000SEP08
205	LG:1096907.1:2000SEP08	481	LG:1096907.1.orf1:2000SEP08
206	LG:1323741.1:2000SEP08	482	LG:1323741.1.orf2:2000SEP08
207	LG:1098372.1:2000SEP08	483	LG:1098372.1.orf1:2000SEP08
208	LG:1006783.1:2000SEP08	484	LG:1006783.1.orf1:2000SEP08
209	LG:1097562.1:2000SEP08	485	LG:1097562.1.orf2:2000SEP08
210	LG:998868.1:2000SEP08	486	LG:998868.1.orf2:2000SEP08
211	LG:1063383.1:2000SEP08	487	LG:1063383.1.orf1:2000SEP08
212	LG:1400567.1:2000SEP08	488	LG:1400567.1.orf2:2000SEP08
213	LI:449404.1:2000SEP08	489	LI:449404.1.orf1:2000SEP08
214	LI:449941.2:2000SEP08	490	LI:449941.2.orf1:2000SEP08
215	LI:450229.1:2000SEP08	491	LI:450229.1.orf1:2000SEP08
216	LI:450399.3:2000SEP08	492	LI:450399.3.orf3:2000SEP08
217	LI:455771.1:2000SEP08	493	LI:455771.1.orf3:2000SEP08
218	LI:720459.1:2000SEP08	494	LI:720459.1.orf1:2000SEP08
219	LI:723156.1:2000SEP08	495	LI:723156.1.orf3:2000SEP08
220	LI:728055.1:2000SEP08	496	LI:728055.1.orf3:2000SEP08
221	LI:1020789.1:2000SEP08	497	LI:1020789.1.orf1:2000SEP08
222	LI:1071728.1:2000SEP08	498	LI:1071728.1.orf1:2000SEP08
223	LI:1084329.1:2000SEP08	499	LI:1084329.1.orf2:2000SEP08
224	LI:246422.1:2000SEP08	500	LI:246422.1.orf1:2000SEP08
225	LI:1086066.1:2000SEP08	501	LI:1086066.1.orf2:2000SEP08
226	LI:223142.1:2000SEP08	502	LI:223142.1.orf3:2000SEP08
227	LI:885368.1:2000SEP08	503	LI:885368.1.orf2:2000SEP08
228	LI:481782.1:2000SEP08	504	LI:481782.1.orf2:2000SEP08
229	LI:1093813.1:2000SEP08	505	LI:1093813.1.orf1:2000SEP08
230	LI:449413.2:2000SEP08	506	LI:449413.2.orf3:2000SEP08
231	LI:450105.1:2000SEP08	507	LI:450105.1.orf1:2000SEP08
231	LI:450105.1:2000SEP08	508	LI:450105.1.orf3:2000SEP08
232	LI:814285.1:2000SEP08	509	LI:814285.1.orf1:2000SEP08
233	LI:1142855.1:2000SEP08	510	LI:1142855.1.orf2:2000SEP08
234	LI:817330.1:2000SEP08	511	LI:817330.1.orf3:2000SEP08
235	LI:817845.1:2000SEP08	512	LI:817845.1.orf2:2000SEP08
236	LI:460809.1:2000SEP08	513	LI:460809.1.orf1:2000SEP08
237	LI:815874.1:2000SEP08	514	LI:815874.1.orf3:2000SEP08
238	LI:255713.1:2000SEP08	515	LI:255713.1.orf2:2000SEP08
239	LI:035973.1:2000SEP08	516	LI:035973.1.orf1:2000SEP08
240	LI:1138110.1:2000SEP08	517	LI:1138110.1.orf3:2000SEP08
241	LI:2049074.1:2000SEP08	518	LI:2049074.1.orf2:2000SEP08
242	LI:1092460.1:2000SEP08	519	LI:1092460.1.orf1:2000SEP08
243	LI:399421.1:2000SEP08	520	LI:399421.1.orf3:2000SEP08

TABLE 1

SEQ ID NO:	Template ID	SEQ ID NO:	ORF ID
244	LI:816655.2:2000SEP08	521	LI:816655.2.orf2:2000SEP08
245	LG:414732.1:2000SEP08	522	LG:414732.1.orf1:2000SEP08
246	LG:1140250.1:2000SEP08	523	LG:1140250.1.orf3:2000SEP08
247	LG:174022.1:2000SEP08	524	LG:174022.1.orf1:2000SEP08
248	LI:002811.1:2000SEP08	525	LI:002811.1.orf1:2000SEP08
249	LI:414732.2:2000SEP08	526	LI:414732.2.orf1:2000SEP08
250	LI:1019920.1:2000SEP08	527	LI:1019920.1.orf1:2000SEP08
251	LI:1038336.1:2000SEP08	528	LI:1038336.1.orf2:2000SEP08
252	LI:1177772.11:2000SEP08	529	LI:1177772.11.orf1:2000SEP08
253	LI:205642.2:2000SEP08	530	LI:205642.2.orf3:2000SEP08
254	LG:449685.1:2000SEP08	531	LG:449685.1.orf3:2000SEP08
255	LG:453922.1:2000SEP08	532	LG:453922.1.orf1:2000SEP08
256	LG:476342.3:2000SEP08	533	LG:476342.3.orf2:2000SEP08
256	LG:476342.3:2000SEP08	534	LG:476342.3.orf3:2000SEP08
257	LI:336801.1:2000SEP08	535	LI:336801.1.orf1:2000SEP08
258	LI:449685.1:2000SEP08	536	LI:449685.1.orf2:2000SEP08
259	LI:476342.1:2000SEP08	537	LI:476342.1.orf1:2000SEP08
260	LI:1072804.1:2000SEP08	538	LI:1072804.1.orf2:2000SEP08
261	LI:455450.1:2000SEP08	539	LI:455450.1.orf1:2000SEP08
262	LI:1073699.1:2000SEP08	540	LI:1073699.1.orf1:2000SEP08
263	LI:1013729.1:2000SEP08	541	LI:1013729.1.orf3:2000SEP08
264	LI:2050322.2:2000SEP08	542	LI:2050322.2.orf2:2000SEP08
265	LI:891327.1:2000SEP08	543	LI:891327.1.orf3:2000SEP08
266	LI:2053076.1:2000SEP08	544	LI:2053076.1.orf3:2000SEP08
267	LG:220085.1:2000SEP08	545	LG:220085.1.orf3:2000SEP08
268	LG:406709.1:2000SEP08	546	LG:406709.1.orf3:2000SEP08
269	LG:347863.9:2000SEP08	547	LG:347863.9.orf3:2000SEP08
270	LI:1073027.1:2000SEP08	548	LI:1073027.1.orf1:2000SEP08
271	LI:347635.1:2000SEP08	549	LI:347635.1.orf2:2000SEP08
272	LI:013685.1:2000SEP08	550	LI:013685.1.orf2:2000SEP08
273	LI:406709.1:2000SEP08	551	LI:406709.1.orf3:2000SEP08
274	LI:2052938.1:2000SEP08	552	LI:2052938.1.orf1:2000SEP08
275	LI:213208.1:2000SEP08	553	LI:213208.1.orf3:2000SEP08

Table 2

SEQ ID NO:	Template ID	GI Number	Probability	Score	Annotation
1	LG:405741.3:2000SEP08	g10439273	1.00E-43		Homo sapiens cDNA: FLJ22761 fs, clone KAJA0893.
2	LG:337194.1:2000SEP08	g10439591	0		Homo sapiens cDNA: FLJ23031 fs, clone LNG01932.
3	LG:017108.4:2000SEP08	g7687936	2.00E-18		(ff) (Leishmania major) possible adenylate kinase
4	LG:372569.5:2000SEP08	g3273307	0		(ff) (Rattus norvegicus) Lysophospholipase
5	LG:968765.1:2000SEP08	g1710247	1.00E-154		Human protein disulfide isomerase-related protein P5 mRNA, partial cds.
6	LG:255999.16:2000SEP08	g8132761	3.00E-17		Homo sapiens glutathione transferase omega (GSTO1) mRNA, complete cds.
7	LG:977820.9:2000SEP08	g14017908	0		Homo sapiens mRNA for KIAA1846 protein, partial cds.
8	LI:1071608.1:2000SEP08	g13529277	6.00E-96		Homo sapiens, aldo-keto reductase family 1, member A1 (aldehyde reductase), clone MGC:12529, mRNA, complete cds.
9	LI:1074023.1:2000SEP08	g1124877	2.00E-99		Macaca mulatta GST-pi enzyme mRNA, complete cds.
10	LI:453570.1:2000SEP08	g219663	1.00E-11		Human mRNA for lactoyl glutathione lyase.
11	LI:072072.1:2000SEP08	g10434968	1.00E-102		Homo sapiens cDNA FLJ13105 fs, clone NT2RP3002351, weakly similar to Human mRNA for NAD-dependent methylene tetrahydrofolate dehydrogenase cyclohydrolase (EC 1.5.1.15).
12	LI:148565.4:2000SEP08	g1185554	9.00E-33		(ff) (Zea mays) glyceraldehyde-3-phosphate dehydrogenase
13	LI:368626.4:2000SEP08	g10441003	0		Homo sapiens epidermal lipoxigenase (ALOXE3) mRNA, complete cds.
14	LI:346123.1:2000SEP08	g546735	1.00E-27		(ff) (Stellaria longipes) triose phosphate isomerase, TPI=5.3.1.1 (Stellaria longipes, Goldie, Peptide, 257 aa)
15	LI:335795.11:2000SEP08	g339679	3.00E-09		Human threonyl-HRNA synthetase mRNA, complete cds.
16	LI:246023.2:2000SEP08	g10434527	0		Homo sapiens cDNA FLJ12816 fs, clone NT2RP2002609, weakly similar to 2-HYDROXYMUCONIC SEMIALDEHYDE HYDROLASE (EC 3.1.1.-).
17	LG:1100661.1:2000SEP08	g189410	1.00E-105		Human oxytocin mRNA, complete cds.
18	LG:475856.1:2000SEP08	g13623353	2.00E-25		Homo sapiens, Similar to zinc finger protein 136 (clone pHZ-20), clone MGC:10647, mRNA, complete cds.
19	LG:1015343.1:2000SEP08	g342298	1.00E-117		M. fascicularis somatostatin I mRNA, complete cds.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
20	LG:1400575.1:2000SEP08	g13623353	4.00E-55	Homo sapiens, Similar to zinc finger protein 136 (clone pHZ-20), clone MGC:10647, mRNA, complete cds.
21	LG:1080545.1:2000SEP08	g12052982	1.00E-88	Homo sapiens mRNA; cDNA DKFZp4341610 (from clone DKFZp4341610); complete cds.
22	LG:213947.1:2000SEP08	g7262613	1.00E-12	(f)(Homo sapiens) candidate taste receptor T2R7
23	LI:720641.1:2000SEP08	g1419016	2.00E-69	(f)(Mus musculus) odorant receptor
24	LI:1023894.1:2000SEP08	g7158201	1.00E-43	(f)(Rattus norvegicus) cytokine receptor-like protein CYRL
25	LI:734904.1:2000SEP08	g12276181	0	Homo sapiens FKSG35 (FKSG35) mRNA, complete cds.
26	LI:1178118.1:2000SEP08	g7239175	2.00E-59	Homo sapiens vanilloid receptor gene, partial sequence; CARL and CTNS genes, complete cds; TIP1 gene, partial cds; P2X5b and P2X5a genes, complete cds; and HUMINAE gene, partial (f)(Homo sapiens) candidate taste receptor T2R7
27	LI:213947.1:2000SEP08	g7262613	8.00E-13	H.sapiens DNA for cGMP phosphodiesterase (exons 4-22).
28	LG:407304.1:2000SEP08	g29924	3.00E-12	Homo sapiens tumor endothelial marker 2 (TEM2) mRNA, complete cds.
29	LG:337358.1:2000SEP08	g9857401	0	Human phospholipase A2 mRNA, complete cds.
30	LG:986090.1:2000SEP08	g189952	8.00E-10	Homo sapiens calcium- and diacylglycerol-regulated guanine nucleotide exchange factor 1 (CalDAG-GEF) mRNA, complete (5' Incom)(Homo sapiens) unknown
31	LG:123250.1:2000SEP08	g4225847	3.00E-07	Homo sapiens mRNA for FLJ00004 protein, partial cds.
32	LG:1028774.2:2000SEP08	g4885696	0	(f)(Homo sapiens) GTPase-activating protein 6 isoform 4
33	LG:338927.6:2000SEP08	g7209308	5.00E-15	Human small GTP-binding protein rab30.
34	LG:332944.2:2000SEP08	g5823454	1.00E-75	(f)(Glycine max) Glycine max calcium dependent protein kinase mRNA
35	LI:347174.5:2000SEP08	g1457955	4.00E-07	Human ADP-ribosylation factor 1 mRNA, complete cds.
36	LI:477070.1:2000SEP08	g169931	5.00E-80	Homo sapiens phospholipase C-beta-2 mRNA, complete cds.
37	LI:723144.1:2000SEP08	g178163	1.00E-44	Homo sapiens phospholipase C-beta-2 mRNA, complete cds.
38	LI:1007188.1:2000SEP08	g190039	5.00E-10	Homo sapiens mRNA for G-protein gamma 7, complete cds.
39	LI:1024412.1:2000SEP08	g3149953	5.00E-31	Homo sapiens calneuron 1 (CALN1) mRNA, complete cds.
40	LI:284797.3:2000SEP08	g13183337	0	Homo sapiens chromosome 21 segment HS21C001.
41	LI:1092901.1:2000SEP08	g7717240	1.00E-41	Homo sapiens phospholipase C beta 1 mRNA, complete cds.
42	LI:228930.1:2000SEP08	g9438228	0	



Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
43	LI:722913.1:2000SEP08	g3450893	1.00E-107	(f) (Avena fatua) ras-like small monomeric GTP-binding protein
44	LG:457478.1:2000SEP08	g13279025	1.00E-108	Homo sapiens. Similar to chromobox homolog 2 (Drosophila Pc class), clone MGC:10561, mRNA, complete cds.
45	LG:358719.1:2000SEP08	g9652377	0	Homo sapiens chromosome 13, partial sequence and human adenovirus type 5 E1A nucleoprotein gene, partial cds.
46	LG:105160.5:2000SEP08	g286024	4.00E-06	Human mRNA for transcription factor, E4TF1-53, complete cds.
47	LG:400705.1:2000SEP08	g488286	2.00E-48	Human basic helix-loop-helix transcription factor mRNA, complete cds.
48	LG:221977.1:2000SEP08	g188909	2.00E-89	Human t(8;17) BCL3/myc gene translocation.
49	LG:898771.1:2000SEP08	g14042881	2.00E-52	Homo sapiens cDNA FLJ14977 fis, clone THYRO1001809, highly similar to MYOCYTE NUCLEAR FACTOR.
50	LI:457478.1:2000SEP08	g13279025	1.00E-108	Homo sapiens. Similar to chromobox homolog 2 (Drosophila Pc class), clone MGC:10561, mRNA, complete cds.
51	LI:125140.1:2000SEP08	g4128144	3.00E-32	Homo sapiens RP58 gene, complete CDS.
52	LI:021095.2:2000SEP08	g37057	1.00E-70	Human mRNA for general transcription factor IIB.
53	LI:888730.1:2000SEP08	g10439413	1.00E-102	Homo sapiens cDNA: FLJ22881 fis, clone KAT03571, highly similar to HUMFERL Human ferritin L chain mRNA.
54	LI:358719.1:2000SEP08	g9652377	1.00E-142	Homo sapiens chromosome 13, partial sequence and human adenovirus type 5 E1A nucleoprotein gene, partial cds.
55	LI:351342.3:2000SEP08	g14042881	0	Homo sapiens cDNA FLJ14977 fis, clone THYRO1001809, highly similar to MYOCYTE NUCLEAR FACTOR.
56	LI:256099.2:2000SEP08	g1561727	1.00E-148	Human transcription factor RTEF-1 (RTEF1) mRNA, complete cds.
57	LI:205199.1:2000SEP08	g2275152	1.00E-116	Homo sapiens DNA-binding protein mRNA, complete cds.
58	LG:980769.1:2000SEP08	g14330447	0	Homo sapiens mRNA for zinc finger protein RINZF (RINZF gene).
59	LG:332474.3:2000SEP08	g9280077	2.00E-60	Macaca fascicularis brain cDNA, clone:Qccc-14453.
60	LG:1087707.1:2000SEP08	g347905	3.00E-41	Human zinc finger protein (ZNF141) mRNA, complete cds.
61	LG:415349.1:2000SEP08	g2244657	1.00E-43	H.sapiens DNA fragment located on chromosome Xq24 containing CpG islands.
62	LG:132420.2:2000SEP08	g1514586	1.00E-28	H.sapiens pseudogene for kruppel-like protein (ZNF75b).

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
63	LG:394201.1:2000SEP08	g2306771	2.00E-10	Human zinc finger protein (LD5-1) gene, exons 4, 5 and 6, and complete cds.
64	LG:1060884.1:2000SEP08	g340443	1.00E-05	Human zinc finger protein 41 (ZNF41) gene, 3' end.
65	LG:242191.1:2000SEP08	g13097776	3.00E-14	Homo sapiens, clone IMAGE:3355405, mRNA, partial cds.
66	LG:1063762.3:2000SEP08	g7022522	2.00E-27	Homo sapiens cDNA FLJ10469 fis, clone NT2RP2000008, weakly similar to ZINC FINGER PROTEIN 84.
67	LG:1100856.1:2000SEP08	g2190183	1.00E-81	Homo sapiens mRNA for zinc finger protein, complete cds.
68	LG:979390.2:2000SEP08	g12804322	5.00E-09	Homo sapiens, clone MGC:4054, mRNA, complete cds.
69	LG:1400447.1:2000SEP08	g487782	4.00E-23	Human zinc finger protein ZNF133.
70	LG:1400562.1:2000SEP08	g454818	1.00E-135	Human Krueppel-related DNA-binding protein (PF4) mRNA, 5'
71	LG:1076130.1:2000SEP08	g5262556	1.00E-14	Homo sapiens mRNA; cDNA DKFZp569D2231 (from clone DKFZp569D2231); partial cds.
72	LG:1064459.1:2000SEP08	g13938260	5.00E-21	Homo sapiens, clone MGC:15514, mRNA, complete cds.
73	LG:1079415.14:2000SEP08	g1049300	3.00E-41	Human KRAB zinc finger protein (ZNF177) mRNA, complete cds.
74	LG:1329431.3:2000SEP08	g4164082	1.00E-155	Homo sapiens zinc finger protein EZNF (EZNF) mRNA, complete cds.
75	LG:1088431.2:2000SEP08	g10439208	1.00E-166	Homo sapiens cDNA: FLJ22713 fis, clone HSI13536.
76	LG:1329462.2:2000SEP08	g10439974	0	Homo sapiens cDNA: FLJ23327 fis, clone HEP12630, highly similar to HSNF37 Homo sapiens ZNF37A mRNA for zinc finger protein.
77	LI:393468.1:2000SEP08	g2618752	1.00E-112	(f)(Taktifugu rubripes) zinc finger protein
78	LI:722577.1:2000SEP08	g5107180	3.00E-31	(f)(Lycopersicon esculentum) small zinc finger-like protein
79	LI:322783.16:2000SEP08	g12053166	0	Homo sapiens mRNA; cDNA DKFZp434O1427 (from clone DKFZp434O1427); complete cds.
80	LI:901355.2:2000SEP08	g8163823	3.00E-79	Homo sapiens krueppel-like zinc finger protein HZF2 mRNA, complete cds.
81	LI:038859.2:2000SEP08	g14330447	0	Homo sapiens mRNA for zinc finger protein RINZF (RINZF gene).
82	LI:1046117.1:2000SEP08	g7023215	8.00E-12	Homo sapiens cDNA FLJ10891 fis, clone NT2RP4002078, weakly similar to ZINC FINGER PROTEIN 91.
83	LI:801015.1:2000SEP08	g13623353	6.00E-42	Homo sapiens, Similar to zinc finger protein 136 (clone pHZ-20), clone MGC:10647, mRNA, complete cds.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
84	LI:1175590.1:2000SEP08	g10432937	0	Homo sapiens cDNA FLJ11637 fs, clone HEMBA1004321, weakly similar to ZINC FINGER PROTEIN 184.
85	LI:1170585.2:2000SEP08	g10439974	2.00E-10	Homo sapiens cDNA: FLJ23327 fs, clone HEP12630, highly similar to HSZNF37 Homo sapiens ZNF37A mRNA for zinc finger protein.
86	LI:1719531.2:2000SEP08	g6650686	7.00E-47	Homo sapiens Y-linked zinc finger protein (ZFY) gene, complete cds.
87	LI:794623.1:2000SEP08	g12804414	3.00E-93	Homo sapiens, Similar to hypothetical protein FLJ10891, clone MGC:925, mRNA, complete cds.
88	LI:1173119.1:2000SEP08	g7959276	2.00E-32	Homo sapiens mRNA for KIAA1508 protein, partial cds.
89	LI:1093285.1:2000SEP08	g10436361	2.00E-62	Homo sapiens cDNA FLJ14012 fs, clone Y79AA1002482, moderately similar to ZINC FINGER PROTEIN 91.
90	LI:1091881.1:2000SEP08	g9802036	1.00E-129	Homo sapiens zinc finger protein SBZF3 mRNA, complete cds.
91	LI:1091617.1:2000SEP08	g5730195	1.00E-21	Homo sapiens partial gene encoding novel Kruppel-type zinc finger, exon 1.
92	LI:1082344.1:2000SEP08	g340473	2.00E-50	Homo sapiens DNA-binding protein (ZNF) gene, partial cds.
93	LI:1166249.1:2000SEP08	g186773	1.00E-100	Human Kruppel related zinc finger protein (HTF10) mRNA, complete cds.
94	LI:799675.1:2000SEP08	g10434780	1.00E-32	Homo sapiens cDNA FLJ12985 fs, clone NT2RP3000050, moderately similar to ZINC FINGER PROTEIN 91.
95	LI:1178899.1:2000SEP08	g10439974	0	Homo sapiens cDNA: FLJ23327 fs, clone HEP12630, highly similar to HSZNF37 Homo sapiens ZNF37A mRNA for zinc finger protein.
96	LI:1169241.1:2000SEP08	g9502201	0	Homo sapiens endotheial zinc finger protein induced by tumor necrosis factor alpha (EZFI) mRNA, complete cds.
97	LI:1180090.1:2000SEP08	g10433741	1.00E-24	Homo sapiens cDNA FLJ12298 fs, clone MAMMA1001837, weakly similar to ZINC FINGER PROTEIN 29.
98	LI:2049322.1:2000SEP08	g10436361	3.00E-76	Homo sapiens cDNA FLJ14012 fs, clone Y79AA1002482, moderately similar to ZINC FINGER PROTEIN 91.
99	LI:809074.1:2000SEP08	g10434194	1.00E-175	Homo sapiens cDNA FLJ12606 fs, clone NT2RM4001483, moderately similar to ZINC FINGER PROTEIN 136.
100	LI:805158.1:2000SEP08	g10437946	2.00E-53	Homo sapiens cDNA: FLJ21782 fs, clone HEP00266, highly similar to AF118063 Homo sapiens PRO1400 mRNA.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
101	LI:1172697.1:2000SEP08	g10435737	0	Homo sapiens cDNA FLJ13659 fis, clone PLACE1011576, moderately similar to Human Kruppel related zinc finger protein (HIF10) mRNA.
102	LI:1174107.2:2000SEP08	g13623586	0	Homo sapiens, Similar to zinc finger protein 254, clone MGC:10544, mRNA, complete cds.
103	LI:1177434.2:2000SEP08	g9968289	1.00E-43	Homo sapiens mRNA for zinc finger protein (ZNF304 gene).
104	LI:1184255.1:2000SEP08	g10436605	2.00E-25	Homo sapiens cDNA FLJ14206 fis, clone NT2RP3003157.
105	LI:1164555.1:2000SEP08	g13938350	9.00E-64	Homo sapiens, Similar to zinc finger protein 268, clone IMAGE:3352268, mRNA, partial cds.
106	LI:238666.4:2000SEP08	g7022522	2.00E-27	Homo sapiens cDNA FLJ10469 fis, clone NT2RP2000008, weakly similar to ZINC FINGER PROTEIN 84.
107	LI:1166752.1:2000SEP08	g10434046	9.00E-09	Homo sapiens cDNA FLJ12515 fis, clone NT2RM2001771, moderately similar to ZINC FINGER PROTEIN 135.
108	LI:2049654.1:2000SEP08	g487782	4.00E-23	Human zinc finger protein ZNF133.
109	LI:242665.2:2000SEP08	g6650686	5.00E-16	Homo sapiens Y-linked zinc finger protein (ZFY) gene, complete
110	LI:208637.1:2000SEP08	g2244657	5.00E-43	H.sapiens DNA fragment located on chromosome Xq24 containing CpG islands.
111	LI:2051808.1:2000SEP08	g1399027	1.00E-175	Human cysteine-rich protein 2 (hCRP2) mRNA, complete cds.
112	LI:1175136.1:2000SEP08	g12803656	1.00E-47	Homo sapiens, Similar to zinc finger protein homologous to mouse Zfp93, clone MGC:3594, mRNA, complete cds.
113	LI:1177337.1:2000SEP08	g10435737	1.00E-147	Homo sapiens cDNA FLJ13659 fis, clone PLACE1011576, moderately similar to Human Kruppel related zinc finger protein (HIF10) mRNA.
114	LI:1165056.1:2000SEP08	g10436361	1.00E-65	Homo sapiens cDNA FLJ14012 fis, clone Y79AA1002482, moderately similar to ZINC FINGER PROTEIN 91.
115	LI:1175250.1:2000SEP08	g10434194	0	Homo sapiens cDNA FLJ12606 fis, clone NT2RM4001483, moderately similar to ZINC FINGER PROTEIN 136.
116	LI:1183192.1:2000SEP08	g13938260	8.00E-21	Homo sapiens, clone MGC:15514, mRNA, complete cds.
117	LI:1183325.1:2000SEP08	g10435640	0	Homo sapiens cDNA FLJ13590 fis, clone PLACE1009398, moderately similar to ZINC FINGER PROTEIN 135.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
118	LI:1178269.2:2000SEP08	g4753759	5.00E-16	Homo sapiens OZF gene exon 1.
119	LI:813422.1:2000SEP08	g10439208	2.00E-83	Homo sapiens cDNA: FLJ22713 fls, clone HS13536.
120	LI:1093049.6:2000SEP08	g881563	4.00E-13	Human zinc finger containing protein ZNF157 (ZNF157) mRNA, complete cds.
121	LI:202192.4:2000SEP08	g1399185	7.00E-10	(fl)(Gallus gallus) zinc finger 5 protein
122	LG:1041854.1:2000SEP08	g12803168	2.00E-69	Homo sapiens, ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit, clone MGC:8347, mRNA, complete cds.
123	LG:1100502.1:2000SEP08	g3659900	2.00E-92	Homo sapiens F1F0-type ATP synthase subunit g mRNA, complete cds.
124	LI:726414.1:2000SEP08	g3319340	4.00E-68	(fl)(Arabidopsis thaliana) contains similarity to E. coli cation transport protein ChaC (GB:D90756)
125	LI:400517.4:2000SEP08	g10439793	0	Homo sapiens cDNA: FLJ23188 fls, clone LNG12038.
126	LI:1078917.1:2000SEP08	g339468	9.00E-29	Human transferrin mRNA, 3' end.
127	LI:1012560.1:2000SEP08	g9957541	0	Homo sapiens connexin 59 (CX59) gene, complete cds.
128	LI:427997.4:2000SEP08	g6996442	1.00E-58	(fl)(Homo sapiens) CTL1 protein
129	LI:197899.1:2000SEP08	g14042128	0	Homo sapiens cDNA FLJ14541 fls, clone NT2RM2001499, moderately similar to LOW-AFFINITY CATIONIC AMINO ACID TRANSPORTER-2.
130	LG:334199.1:2000SEP08	g14017846	7.00E-39	Homo sapiens mRNA for KIAA1815 protein, partial cds.
131	LG:334345.1:2000SEP08	g14336735	4.00E-08	Homo sapiens 16p13.3 sequence section 5 of 8.
132	LG:228092.1:2000SEP08	g13529112	0	Homo sapiens, clone IMAGE:3930327, mRNA, partial cds.
133	LG:098580.1:2000SEP08	g4096351	3.00E-09	Human apoptotic cysteine protease Mch1/TX isoform delta (mch1/Tx) mRNA, complete cds.
134	LG:969572.1:2000SEP08	g177889	2.00E-07	Human alpha-2-thiol proteinase inhibitor mRNA, complete coding sequence.
135	LG:196958.1:2000SEP08	g1182066	1.00E-07	Human tryptase mRNA, complete cds.
136	LG:1087811.1:2000SEP08	g2565302	1.00E-130	Macaca mulatta cyclophilin A mRNA, complete cds.
137	LG:1327885.1:2000SEP08	g13182746	0	Homo sapiens microsomal signal peptidase subunit mRNA, complete cds.
138	LI:449393.1:2000SEP08	g14348899	2.00E-07	Homo sapiens heat shock protein mRNA, complete cds.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
139	LI:897616.1:2000SEP08	g337369	1.00E-103	Human rapamycin- and FK506-binding protein, complete cds.
140	LI:736860.1:2000SEP08	g475922	1.00E-24	(f) (Zea mays) proteinase inhibitor
141	LI:027066.6:2000SEP08	g10432866	2.00E-95	Homo sapiens cDNA FLJ11583 fs, clone HEMBA1003680, weakly similar to PUTATIVE AMINOPEPTIDASE ZK353.6 IN CHROMOSOME III (EC 3.4.11.-).
142	LI:1074263.1:2000SEP08	g2281120	6.00E-38	Saimiri sciureus cystatin C mRNA, complete cds.
143	LI:334345.1:2000SEP08	g14336735	4.00E-08	Homo sapiens 16p13.3 sequence section 5 of 8.
144	LI:1093914.1:2000SEP08	g5926696	1.00E-175	Homo sapiens genomic DNA, chromosome 6p21.3, HLA Class I region, section 8/20.
145	LI:1188168.1:2000SEP08	g10503945	0	Homo sapiens calpain-like protease CAPN10e mRNA, complete cds.
146	LI:1065168.1:2000SEP08	g12653472	6.00E-90	Homo sapiens, proteasome (prosome, macropain) subunit, beta type, 1, clone MGC:8505, mRNA, complete cds.
147	LI:1180418.1:2000SEP08	g2565300	1.00E-130	Cercopithecus aethiops cyclophilin A mRNA, complete cds.
148	LG:232648.1:2000SEP08	g14336723	0	Homo sapiens 16p13.3 sequence section 4 of 8.
149	LG:1078420.1:2000SEP08	g1263080	1.00E-145	Human mariner1 transposase gene, complete consensus
150	LG:1397599.1:2000SEP08	g2104909	5.00E-97	Human endogenous retrovirus H D1 leader region/integrase-derived ORF1, ORF2, and putative envelope protein mRNA.
151	LG:1397655.2:2000SEP08	g2104909	1.00E-101	Human endogenous retrovirus H D1 leader region/integrase-derived ORF1, ORF2, and putative envelope protein mRNA.
152	LG:241055.1:2000SEP08	g1263080	1.00E-173	Human mariner1 transposase gene, complete consensus
153	LG:1101065.1:2000SEP08	g2226003	9.00E-44	Human Tigger1 transposable element, complete consensus sequence.
154	LG:475629.1:2000SEP08	g4185140	1.00E-39	(f) (Arabidopsis thaliana) putative small nuclear
155	LI:348991.1:2000SEP08	g31394	3.00E-80	Human humFib mRNA for fibrillarin.
156	LI:475629.1:2000SEP08	g4185140	1.00E-33	(f) (Arabidopsis thaliana) putative small nuclear
157	LI:261331.1:2000SEP08	g1263080	3.00E-88	Human mariner1 transposase gene, complete consensus
158	LI:815686.1:2000SEP08	g1698454	1.00E-150	Human mariner2 transposable element, complete consensus sequence.
159	LI:1167327.2:2000SEP08	g2104909	3.00E-83	Human endogenous retrovirus H D1 leader region/integrase-derived ORF1, ORF2, and putative envelope protein mRNA.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
160	LI:758009.3:2000SEP08	g9650710	0	Homo sapiens mRNA for HEF like Protein (HEFL gene).
161	LG:331593.1:2000SEP08	g4106980	6.00E-11	(ff) (Homo sapiens) immunoglobulin-like transcript 10 protein
162	LI:1094174.1:2000SEP08	g3273727	4.00E-60	Homo sapiens MHC class 1 region.
163	LI:814362.1:2000SEP08	g261239	1.00E-149	Immunoglobulin M light chain V region=anti-Ipld A antibody (human, hybridoma cell line HR78, mRNA Partial, 460 nt).
164	LI:219542.1:2000SEP08	g30150	1.00E-19	H.sapiens coxIIb mRNA for cytochrome c oxidase subunit VIIb.
165	LI:726197.1:2000SEP08	g2114207	1.00E-48	(ff) (Oryza sativa) glutaredoxin
166	LI:1075314.1:2000SEP08	g2588778	1.00E-100	Homo sapiens mRNA for cytochrome b large subunit of complex II, complete cds.
167	LI:437883.1:2000SEP08	g986883	1.00E-146	Human nuclear-encoded mitochondrial NADH-ubiquinone reductase 24Kd subunit mRNA, complete cds.
168	LG:336265.1:2000SEP08	g159649	2.00E-43	(ff) (Ascaris suum) putative
169	LG:407788.2:2000SEP08	g12052773	0	Homo sapiens mRNA; cDNA DKFZp564B052 (from clone DKFZp564B052); complete cds.
170	LG:1326925.1:2000SEP08	g2853300	1.00E-135	Homo sapiens mucin (MUC3) mRNA, partial cds.
171	LI:332655.2:2000SEP08	g6996452	3.00E-07	Homo sapiens SPP2 gene for secreted phosphoprotein 24 precursor, exons 1-8.
172	LI:1184621.4:2000SEP08	g12052773	3.00E-62	Homo sapiens mRNA; cDNA DKFZp564B052 (from clone DKFZp564B052); complete cds.
173	LI:2051386.1:2000SEP08	g3228236	4.00E-49	Homo sapiens UHS KerB gene.
174	LG:362757.1:2000SEP08	g1419370	7.00E-74	(ff) (Zea mays) actin depolymerizing factor
175	LG:406770.1:2000SEP08	g508483	3.00E-48	Homo sapiens GATA-4 mRNA, complete cds.
176	LG:1094640.1:2000SEP08	g13249136	0	Homo sapiens chromosome 2 unknown sequence.
177	LG:001929.1:2000SEP08	g908802	1.00E-44	Homo sapiens keratin 6 isoform K6e (KRT6E) mRNA, complete
178	LI:401322.1:2000SEP08	g13623540	6.00E-23	Homo sapiens, tubulin alpha 1, clone MGC:12832, mRNA, complete cds.
179	LI:208748.1:2000SEP08	g10433083	0	Homo sapiens cDNA FLJ11756 fis, clone HEMBA1005595, weakly similar to DYNEIN HEAVY CHAIN, CYTOSOLIC.
180	LI:407242.1:2000SEP08	g2282582	4.00E-71	(ff) (Mus musculus) actin-binding protein
181	LI:403409.1:2000SEP08	g8896163	0	Homo sapiens kinesin-like protein GAKIN mRNA, complete cds.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
182	LI:450798.1:2000SEP08	g12804324	2.00E-15	Homo sapiens, clone IMAGE:2823044, mRNA, partial cds.
183	LI:410317.1:2000SEP08	g186699	1.00E-53	Human 56k cytoskeletal type II keratin mRNA.
184	LI:340268.1:2000SEP08	g12653054	4.00E-68	Homo sapiens, actin, gamma 1, clone MGC:8332, mRNA, complete cds.
185	LI:2051671.1:2000SEP08	g63805	1.00E-38	(ff)(Gallus gallus) tensin
186	LG:998844.1:2000SEP08	g7259234	3.00E-22	(ff)(Mus musculus) contains transmembrane (TM) region
187	LG:1043787.1:2000SEP08	g311699	1.00E-156	H.sapiens GPx-4 mRNA for phospholipid hydroperoxide glutathione peroxidase.
188	LG:1098931.16:2000SEP08	g2138329	4.00E-48	Human acetyl-CoA carboxylase (ACC2) mRNA, complete cds.
189	LG:199423.2:2000SEP08	g4520343	1.00E-12	Homo sapiens mRNA for N-copine, complete cds.
190	LI:1075297.1:2000SEP08	g13543530	1.00E-36	Homo sapiens, microsomal glutathione S-transferase 1, clone MGC:14525, mRNA, complete cds.
191	LI:1043321.1:2000SEP08	g546517	3.00E-71	stearoyl-CoA desaturase (human, adipose tissue, mRNA Partial, 712 nt).
192	LI:297070.1:2000SEP08	g307297	8.00E-26	Human I beta 1-6 N-acetylglucosaminyltransferase mRNA, complete cds.
193	LI:1085041.1:2000SEP08	g13543567	8.00E-34	Homo sapiens, prostaglandin D2 synthase (21kD, brain), clone MGC:14559, mRNA, complete cds.
194	LI:1071544.1:2000SEP08	g35069	1.00E-161	H.sapiens RNA for nm23-H2 gene.
195	LI:2052480.1:2000SEP08	g10439273	0	Homo sapiens cDNA: FLJ22761 fs, clone KAI0893.
196	LG:450105.1:2000SEP08	g414348	8.00E-09	Human homolog of yeast ribosomal protein S28, complete cds.
197	LG:450581.1:2000SEP08	g2739219	7.00E-28	(ff)(Hordeum vulgare) rps28
198	LG:450887.1:2000SEP08	g7629994	1.00E-40	(ff)(Arabidopsis thaliana) 60S RIBOSOMAL PROTEIN L36 homolog
199	LG:460809.1:2000SEP08	g36129	5.00E-55	Human mRNA for ribosomal protein L31.
200	LG:452089.1:2000SEP08	g7340874	4.00E-88	(ff)(Oryza sativa) ESTs
				D15590(C0900), D48950(S15542), D22684(C0900) correspond to a region of the predicted gene. ~Similar to Arabidopsis thaliana 60S ribosomal protein L11A (L16A). (P42795)
201	LG:1099416.1:2000SEP08	g292440	1.00E-85	Homo sapiens ribosomal protein L37 mRNA, complete cds.
202	LG:255713.1:2000SEP08	g36129	3.00E-51	Human mRNA for ribosomal protein L31.



Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
203	LG:998903.1:2000SEP08	g5106775	4.00E-69	(f1)(Hordeum vulgare) ribosomal protein S12
204	LG:1119656.1:2000SEP08	g1220360	6.00E-13	Homo sapiens (clone corf-1c15) S29 ribosomal protein mRNA, complete cds.
205	LG:1096907.1:2000SEP08	g292438	8.00E-98	Homo sapiens ribosomal protein L37a (RPL37A) mRNA, complete cds.
206	LG:1323741.1:2000SEP08	g562073	1.00E-127	Human ribosomal protein L35 mRNA, complete cds.
207	LG:1098372.1:2000SEP08	g31061	1.00E-121	Human mRNA for Epstein-Barr virus small RNAs (EBERs) associated protein (EAP).
208	LG:1006783.1:2000SEP08	g550014	1.00E-161	Human ribosomal protein L21 mRNA, complete cds.
209	LG:1097562.1:2000SEP08	g488414	1.00E-148	H.sapiens mRNA for ribosomal protein L30.
210	LG:998868.1:2000SEP08	g483431	1.00E-134	(f1)(Oryza sativa) cyc07
211	LG:1063383.1:2000SEP08	g36131	8.00E-72	Human mRNA for ribosomal protein L32.
212	LG:1400567.1:2000SEP08	g505472	1.00E-107	H.sapiens mRNA for ribosomal protein L31.
213	LI:449404.1:2000SEP08	g13905003	1.00E-29	Homo sapiens, ribosomal protein S14, clone MGC:5275, mRNA, complete cds.
214	LI:449941.2:2000SEP08	g968902	1.00E-106	(f1)(Oryza sativa) ribosomal protein S8
215	LI:450229.1:2000SEP08	g4588906	7.00E-97	(f1)(Secale cereale) ribosomal protein S7
216	LI:450399.3:2000SEP08	g36125	2.00E-10	Human mRNA for ribosomal protein L17.
217	LI:455771.1:2000SEP08	g414348	2.00E-19	Human homolog of yeast ribosomal protein S28, complete cds.
218	LI:720459.1:2000SEP08	g13278716	1.00E-130	Homo sapiens, ribosomal protein L6, clone MGC:1635, mRNA, complete cds.
219	LI:723156.1:2000SEP08	g915313	3.00E-53	(f1)(Nicotiana glutinosa) ribosomal protein L31
220	LI:728055.1:2000SEP08	g14044115	1.00E-140	Homo sapiens, ribosomal protein S16, clone MGC:15283, mRNA, complete cds.
221	LI:1020789.1:2000SEP08	g38422	1.00E-146	Homo sapiens mRNA for ribosomal protein S18.
222	LI:1071728.1:2000SEP08	g13960132	1.00E-139	Homo sapiens, ribosomal protein S20, clone MGC:4151, mRNA, complete cds.
223	LI:1084329.1:2000SEP08	g14043190	1.00E-156	Homo sapiens, clone MGC:15572, mRNA, complete cds.
224	LI:246422.1:2000SEP08	g409069	1.00E-136	Human mRNA for HBp15/L22, complete cds.
225	LI:1086066.1:2000SEP08	g550020	1.00E-170	Human ribosomal protein S5 mRNA, complete cds.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
226	LI:223142.1:2000SEP08	g12653460	2.00E-20	Homo sapiens, ribosomal protein L17, clone MGC:8457, mRNA, complete cds.
227	LI:885368.1:2000SEP08	g2150130	1.00E-66	(f) (Arabidopsis thaliana) cytoplasmic ribosomal protein S15a
228	LI:481782.1:2000SEP08	g4115717	8.00E-06	Chlorocebus aethiops mRNA for ribosomal protein S4X, complete cds.
229	LI:1093813.1:2000SEP08	g562073	1.00E-125	Human ribosomal protein L35 mRNA, complete cds.
230	LI:449413.2:2000SEP08	g414348	2.00E-15	Human homolog of yeast ribosomal protein S28, complete cds.
231	LI:450105.1:2000SEP08	g414348	3.00E-08	Human homolog of yeast ribosomal protein S28, complete cds.
232	LI:814285.1:2000SEP08	g14250761	2.00E-59	Homo sapiens, clone MGC:14308, mRNA, complete cds.
233	LI:1142855.1:2000SEP08	g10439463	1.00E-146	Homo sapiens cDNA: FLJ22926 fis, clone KAT06984, highly similar to HUMPPARP1 Human acidic ribosomal phosphoprotein P1
234	LI:817330.1:2000SEP08	g5106775	8.00E-28	(f) (Hordeum vulgare) ribosomal protein S12
235	LI:817845.1:2000SEP08	g9759463	3.00E-60	(f) (Arabidopsis thaliana) 40S ribosomal protein S19
236	LI:460809.1:2000SEP08	g505472	7.00E-55	H. sapiens mRNA for ribosomal protein L31.
237	LI:815874.1:2000SEP08	g292442	1.00E-144	Homo sapiens ribosomal protein S20 (RPS20) mRNA, complete cds.
238	LI:255713.1:2000SEP08	g36129	4.00E-52	Human mRNA for ribosomal protein L31.
239	LI:935973.1:2000SEP08	g292440	7.00E-66	Homo sapiens ribosomal protein L37 mRNA, complete cds.
240	LI:1138110.1:2000SEP08	g1220360	6.00E-13	Homo sapiens (clone cori-1 cl5) S29 ribosomal protein mRNA, complete cds.
241	LI:2049074.1:2000SEP08	g13938411	3.00E-95	Homo sapiens, ribosomal protein L13, clone MGC:15415, mRNA, complete cds.
242	LI:1092460.1:2000SEP08	g12652698	1.00E-154	Homo sapiens, purine-rich element binding protein B, clone MGC:1947, mRNA, complete cds.
243	LI:399421.1:2000SEP08	g1568556	1.00E-35	H. sapiens H2B/I gene.
244	LI:816655.2:2000SEP08	g10439444	1.00E-127	Homo sapiens cDNA: FLJ22909 fis, clone KAT05694, highly similar to HUMHMG17 Human non-histone chromosomal protein HMG-17 mRNA.
245	LG:414732.1:2000SEP08	g183232	1.00E-140	Human beta-glucuronidase mRNA, complete cds.
246	LG:1140250.1:2000SEP08	g8101070	4.00E-91	Homo sapiens golgin-like protein (GLP) gene, complete cds.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
247	LG:174022.1:2000SEP08	g2906145	1.00E-160	Homo sapiens malate dehydrogenase precursor (MDH) mRNA, nuclear gene encoding mitochondrial protein, complete cds.
248	LI:002811.1:2000SEP08	g5702305	7.00E-66	Homo sapiens vault protein mRNA, complete cds.
249	LI:414732.2:2000SEP08	g183232	1.00E-142	Human beta-glucuronidase mRNA, complete cds.
250	LI:1019920.1:2000SEP08	g4164451	6.00E-71	Homo sapiens NADH-ubiquinone oxidoreductase B14.5A subunit mRNA, nuclear gene encoding mitochondrial protein,
251	LI:1038336.1:2000SEP08	g4164443	2.00E-24	Homo sapiens NADH:ubiquinone oxidoreductase B9 subunit mRNA, nuclear gene encoding mitochondrial protein,
252	LI:1177772.11:2000SEP08	g7211437	1.00E-141	Homo sapiens golgin-67 (GOLGA5) mRNA, complete cds.
253	LI:205642.2:2000SEP08	g4510363	3.00E-71	(f) (Arabidopsis thaliana) putative ubiquitin-conjugating enzyme
254	LG:449685.1:2000SEP08	g7248411	8.00E-38	(f) (Oryza sativa) ESTs C99632(E20954), C99633(E20954) correspond to a region of the predicted gene. ~Similar to Arabidopsis thaliana putative pathogenesis-related protein
255	LG:453922.1:2000SEP08	g3789950	2.00E-55	(f) (Oryza sativa) translation initiation factor
256	LG:476342.3:2000SEP08	g790641	3.00E-25	(f) (Hordeum vulgare) gamma-thionin
257	LI:336801.1:2000SEP08	g5912457	1.00E-86	(f) (Homo sapiens) dJ1068E13.2 (novel protein similar to bovine SCP2 (Sterol Carrier Protein 2) and part of HSD17B4 (hydroxysteroid (17-beta) dehydrogenase 4))
258	LI:449685.1:2000SEP08	g7248411	9.00E-33	(f) (Oryza sativa) ESTs C99632(E20954), C99633(E20954) correspond to a region of the predicted gene. ~Similar to Arabidopsis thaliana putative pathogenesis-related protein
259	LI:476342.1:2000SEP08	g790641	2.00E-25	(f) (Hordeum vulgare) gamma-thionin
260	LI:1072804.1:2000SEP08	g453189	2.00E-58	(f) (Zea mays) acyl carrier protein
261	LI:455450.1:2000SEP08	g4105111	2.00E-43	(f) (Hordeum vulgare) dehydrin 6
262	LI:1073699.1:2000SEP08	g219661	1.00E-43	Human mRNA for growth inhibitory factor.
263	LI:1013729.1:2000SEP08	g182353	6.00E-41	Human fatty acid binding protein homologue (PA-FABP) mRNA, complete cds.
264	LI:2050322.2:2000SEP08	g10439685	1.00E-115	Homo sapiens cDNA: FLJ23107 fis, clone LNG07738.
265	LI:891327.1:2000SEP08	g8570523	0	Homo sapiens genomic DNA, chromosome 1q22-q23, CD1 region, section 3/4.

Table 2

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
266	LI:2053076.1:2000SEP08	g3879684	7.00E-20	(f1)(Caenorhabditis elegans) (Z74042) predicted using Genefinder-Similarity to Haemophilus 3-oxoacyl-(acyl-carrier protein) reductase (SW:FABG_HAEIN), contains similarity to Pfam domain: PF00106 (short chain dehydrogenase), Score=170.5, E-value=9.2e-48, N=1-cDNA EST yk470b2.3 comes from this gene-cDNA EST yk470b2.5 comes from this gene Human endogenous retrovirus type C oncovirus sequence.
267	LG:220085.1:2000SEP08	g325464	4.00E-98	
268	LG:406709.1:2000SEP08	g5106556	8.00E-11	Homo sapiens MILL septin-like fusion protein (MSF) mRNA, complete cds.
269	LG:347863.9:2000SEP08	g6708478	2.00E-22	(3' incm) (Mus musculus) formin-like protein
270	LI:1073027.1:2000SEP08	g14198256	1.00E-116	Homo sapiens, clone MGC:5243, mRNA, complete cds.
271	LI:347635.1:2000SEP08	g11527996	0	Homo sapiens NOTCH2 protein (NOTCH2) mRNA, complete cds.
272	LI:013685.1:2000SEP08	g4868434	3.00E-79	Homo sapiens apoptosis related protein APR-2 mRNA, complete cds.
273	LI:406709.1:2000SEP08	g5106556	8.00E-08	Homo sapiens MILL septin-like fusion protein (MSF) mRNA, complete cds.
274	LI:2052938.1:2000SEP08	g2662080	1.00E-08	Homo sapiens KIAA0400 mRNA, complete cds.
275	LI:213208.1:2000SEP08	g179303	2.00E-12	Human B12 protein mRNA, complete cds.

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
1	LG:405741.3:2000SEP08	204	380	forward 3	hexokinase	Hexokinase	2.10E-30
3	LG:017108.4:2000SEP08	335	574	forward 2	adenylatekinase	Adenylate kinase	2.40E-09
4	LG:372569.5:2000SEP08	1340	1441	forward 2	ank	Ank repeat	9.50E-05
4	LG:372569.5:2000SEP08	202	1125	forward 1	Asparaginase	Asparaginase	4.60E-49
5	LG:968765.1:2000SEP08	182	433	forward 2	thioredoxin	Thioredoxin	5.90E-40
7	LG:977820.9:2000SEP08	11	844	forward 2	AMP-binding	AMP-binding enzyme	1.70E-13
8	LI:1071608.1:2000SEP08	267	539	forward 3	aldo_ket_red	Aldo/keto reductase family	1.90E-43
9	LI:1074023.1:2000SEP08	294	575	forward 3	GST	Glutathione S-transferase, C-terminal	3.30E-19
9	LI:1074023.1:2000SEP08	60	278	forward 3	GST_N	Glutathione S-transferase, N-terminal	9.90E-20
10	LI:453570.1:2000SEP08	186	626	forward 3	Glyoxalase	Glyoxalase/Bleomycin resistance protein/Dioxygenase superfamily	1.10E-36
11	LI:072072.1:2000SEP08	535	729	forward 1	zf-DHHC	DHHC zinc finger domain	2.50E-25
12	LI:148565.4:2000SEP08	449	580	forward 2	gpdh	Glyceraldehyde 3-phosphate dehydrogenase, NAD binding domain	1.90E-14
12	LI:148565.4:2000SEP08	400	453	forward 1	gpdh	Glyceraldehyde 3-phosphate dehydrogenase, NAD binding domain	3.50E-08
13	LI:368626.4:2000SEP08	105	239	forward 3	lipoxigenase	Lipoxigenase	1.20E-13
13	LI:368626.4:2000SEP08	226	462	forward 1	lipoxigenase	Lipoxigenase	2.50E-05
14	LI:346123.1:2000SEP08	99	302	forward 3	TIM	Triosephosphate isomerase	1.20E-08
14	LI:346123.1:2000SEP08	373	492	forward 1	TIM	Triosephosphate isomerase	1.70E-06
15	LI:335795.11:2000SEP08	512	706	forward 2	TGS	TGS domain	1.30E-11
15	LI:335795.11:2000SEP08	1203	2375	forward 3	tRNA-synt_2b	tRNA synthetase class II (G, H, P, S and T)	2.50E-19
16	LI:246023.2:2000SEP08	319	1050	forward 1	abhydrolase	alpha/beta hydrolase fold	8.20E-20
17	LG:1100661.1:2000SEP08	138	371	forward 3	hormone5	Neurohypophyseal hormones, C-terminal	3.00E-45
18	LG:475856.1:2000SEP08	139	207	forward 1	zf-C2H2	Domain	4.30E-05
20	LG:1400575.1:2000SEP08	152	304	forward 2	KRAB	Zinc finger, C2H2 type	2.30E-16
20	LG:1400575.1:2000SEP08	440	508	forward 2	zf-C2H2	KRAB box	1.40E-07
21	LG:1080545.1:2000SEP08	221	289	forward 2	zf-C2H2	Zinc finger, C2H2 type	1.90E-06
23	LI:720641.1:2000SEP08	236	985	forward 2	7tm_1	7 transmembrane receptor (rhodopsin	1.60E-27
25	LI:734904.1:2000SEP08	1197	1256	forward 3	7tm_1	7 transmembrane receptor (rhodopsin	3.90E-06

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hlt	Pfam Description	E-value
25	LI:734904.1:2000SEP08	964	1182	forward 1	7tm_1	7 transmembrane receptor (rhodopsin	4.20E-05
26	LI:1178118.1:2000SEP08	753	821	forward 3	zf-C2H2	Zinc finger, C2H2 type	2.60E-06
29	LG:337358.1:2000SEP08	701	1438	forward 2	ras	Ras family	5.50E-32
30	LG:986090.1:2000SEP08	85	798	forward 1	14-3-3	14-3-3 proteins	1.30E-145
31	LG:123250.1:2000SEP08	812	1315	forward 2	RasGEF	RasGEF domain	1.30E-07
33	LG:338927.6:2000SEP08	968	1219	forward 2	PH	PH domain	3.60E-07
34	LG:332944.2:2000SEP08	700	1185	forward 1	RhoGAP	RhoGAP domain	3.00E-22
35	LI:347174.5:2000SEP08	23	526	forward 2	arf	ADP-ribosylation factor family	2.80E-04
35	LI:347174.5:2000SEP08	35	640	forward 2	ras	Ras family	5.70E-78
36	LI:477070.1:2000SEP08	83	169	forward 2	efhand	EF hand	7.80E-08
37	LI:723144.1:2000SEP08	68	604	forward 2	arf	ADP-ribosylation factor family	3.80E-129
37	LI:723144.1:2000SEP08	119	625	forward 2	ras	Ras family	1.60E-04
38	LI:1007188.1:2000SEP08	295	534	forward 1	PLC-X	Phosphatidylinositol-specific phospholipase C, X domain	2.30E-41
39	LI:1024412.1:2000SEP08	117	281	forward 3	G-gamma	GGL domain	4.50E-36
40	LI:284797.3:2000SEP08	291	377	forward 3	efhand	EF hand	2.50E-08
43	LI:722913.1:2000SEP08	172	744	forward 1	arf	ADP-ribosylation factor family	6.60E-105
44	LG:457478.1:2000SEP08	98	220	forward 2	chromo	'chromo' (CHRromatin Organization Modifier) domain	1.70E-24
45	LG:358719.1:2000SEP08	285	1151	forward 3	Adeno_E1A	Early E1A protein	6.00E-201
47	LG:400705.1:2000SEP08	277	435	forward 1	HLH	Helix-loop-helix DNA-binding domain	1.10E-17
49	LG:898771.1:2000SEP08	79	291	forward 1	FHA	FHA domain	1.50E-14
49	LG:898771.1:2000SEP08	1825	2112	forward 1	Fork_head	Fork head domain	7.10E-65
50	LI:457478.1:2000SEP08	98	220	forward 2	chromo	'chromo' (CHRromatin Organization Modifier) domain	1.70E-24
51	LI:125140.1:2000SEP08	143	469	forward 2	BTB	BTB/POZ domain	8.40E-06
53	LI:888730.1:2000SEP08	337	606	forward 1	ferritin	Ferritin	1.80E-44
53	LI:888730.1:2000SEP08	225	344	forward 3	ferritin	Ferritin	9.00E-17
54	LI:358719.1:2000SEP08	291	1109	forward 3	Adeno_E1A	Early E1A protein	5.60E-14
54	LI:358719.1:2000SEP08	295	1161	forward 1	Adeno_E1A	Early E1A protein	2.60E-10

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit IRF	Pfam Description	E-value
57	LI:2051991.1:2000SEP08	340	588	forward 1		Interferon regulatory factor transcription factor	1.10E-13
59	LG:332474.3:2000SEP08	463	624	forward 1	KRAB	KRAB box	6.30E-22
60	LG:1087707.1:2000SEP08	454	642	forward 1	KRAB	KRAB box	2.80E-40
62	LG:132420.2:2000SEP08	176	463	forward 2	SCAN	SCAN domain	7.60E-47
64	LG:1060884.1:2000SEP08	798	866	forward 3	zf-C2H2	Zinc finger, C2H2 type	4.50E-04
65	LG:242191.1:2000SEP08	821	889	forward 2	zf-C2H2	Zinc finger, C2H2 type	1.20E-06
66	LG:1063762.3:2000SEP08	268	456	forward 1	KRAB	KRAB box	1.30E-46
66	LG:1063762.3:2000SEP08	859	927	forward 1	zf-C2H2	Zinc finger, C2H2 type	2.40E-05
68	LG:979390.2:2000SEP08	85	273	forward 1	KRAB	KRAB box	2.40E-34
69	LG:1400447.1:2000SEP08	386	571	forward 2	KRAB	KRAB box	6.80E-37
70	LG:1400562.1:2000SEP08	162	350	forward 3	KRAB	KRAB box	6.50E-40
71	LG:1076130.1:2000SEP08	564	632	forward 3	zf-C2H2	Zinc finger, C2H2 type	3.20E-07
72	LG:1064459.1:2000SEP08	38	229	forward 2	KRAB	KRAB box	2.60E-29
72	LG:1064459.1:2000SEP08	458	526	forward 2	zf-C2H2	Zinc finger, C2H2 type	2.60E-06
74	LG:1329431.3:2000SEP08	395	583	forward 2	KRAB	KRAB box	8.30E-41
75	LG:1088431.2:2000SEP08	199	363	forward 1	KRAB	KRAB box	5.00E-21
76	LG:1329462.2:2000SEP08	1006	1191	forward 1	KRAB	KRAB box	7.10E-39
77	LI:393468.1:2000SEP08	557	625	forward 2	zf-C2H2	Zinc finger, C2H2 type	1.30E-04
79	LI:322783.16:2000SEP08	478	591	forward 1	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1.80E-06
80	LI:901355.2:2000SEP08	260	448	forward 2	KRAB	KRAB box	2.40E-41
82	LI:1046117.1:2000SEP08	85	231	forward 1	KRAB	KRAB box	3.70E-16
83	LI:801015.1:2000SEP08	22	216	forward 1	KRAB	KRAB box	3.00E-24
84	LI:1175590.1:2000SEP08	225	293	forward 3	zf-C2H2	Zinc finger, C2H2 type	2.90E-05
84	LI:1175590.1:2000SEP08	380	448	forward 2	zf-C2H2	Zinc finger, C2H2 type	4.40E-05
85	LI:1170585.2:2000SEP08	247	435	forward 1	KRAB	KRAB box	1.10E-38
85	LI:1170585.2:2000SEP08	635	703	forward 2	zf-C2H2	Zinc finger, C2H2 type	1.40E-06
87	LI:794623.1:2000SEP08	128	238	forward 2	KRAB	KRAB box	3.60E-15
88	LI:1173119.1:2000SEP08	642	710	forward 3	zf-C2H2	Zinc finger, C2H2 type	9.30E-06
89	LI:1093285.1:2000SEP08	112	180	forward 1	zf-C2H2	Zinc finger, C2H2 type	5.70E-07
90	LI:1091881.1:2000SEP08	133	321	forward 1	KRAB	KRAB box	5.10E-40

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
91	LI:1091617.1:2000SEP08	392	463	forward 2	zf-C2H2	Zinc finger, C2H2 type	9.80E-07
92	LI:1082344.1:2000SEP08	487	555	forward 1	zf-C2H2	Zinc finger, C2H2 type	8.80E-06
93	LI:1166249.1:2000SEP08	99	287	forward 3	KRAB	KRAB box	6.10E-43
94	LI:799675.1:2000SEP08	262	411	forward 1	KRAB	KRAB box	9.50E-15
95	LI:1178899.1:2000SEP08	1052	1237	forward 2	KRAB	KRAB box	7.10E-39
96	LI:1169241.1:2000SEP08	405	473	forward 3	zf-C2H2	Zinc finger, C2H2 type	1.50E-07
97	LI:1180090.1:2000SEP08	277	564	forward 1	SCAN	SCAN domain	9.40E-61
98	LI:2049322.1:2000SEP08	127	315	forward 1	KRAB	KRAB box	2.00E-34
99	LI:809074.1:2000SEP08	176	328	forward 2	KRAB	KRAB box	1.10E-15
100	LI:805158.1:2000SEP08	602	670	forward 2	zf-C2H2	Zinc finger, C2H2 type	5.90E-08
100	LI:805158.1:2000SEP08	126	194	forward 3	zf-C2H2	Zinc finger, C2H2 type	4.90E-06
101	LI:1172697.1:2000SEP08	229	420	forward 1	KRAB	KRAB box	2.00E-44
102	LI:1174107.2:2000SEP08	111	302	forward 3	KRAB	KRAB box	2.00E-43
103	LI:1177434.2:2000SEP08	279	446	forward 3	KRAB	KRAB box	4.50E-16
103	LI:1177434.2:2000SEP08	851	919	forward 2	zf-C2H2	Zinc finger, C2H2 type	8.90E-08
104	LI:1184255.1:2000SEP08	65	133	forward 2	zf-C2H2	Zinc finger, C2H2 type	3.70E-06
105	LI:1164555.1:2000SEP08	288	356	forward 3	zf-C2H2	Zinc finger, C2H2 type	2.40E-05
106	LI:238666.4:2000SEP08	268	456	forward 1	KRAB	KRAB box	1.30E-46
106	LI:238666.4:2000SEP08	859	927	forward 1	zf-C2H2	Zinc finger, C2H2 type	2.40E-05
108	LI:2049654.1:2000SEP08	343	528	forward 1	KRAB	KRAB box	6.80E-37
110	LI:208637.1:2000SEP08	187	516	forward 1	BTB	BTB/POZ domain	1.40E-23
111	LI:2051808.1:2000SEP08	385	558	forward 1	LIM	LIM domain	5.50E-17
113	LI:1177337.1:2000SEP08	319	387	forward 1	zf-C2H2	Zinc finger, C2H2 type	1.70E-07
113	LI:1177337.1:2000SEP08	476	547	forward 2	zf-C2H2	Zinc finger, C2H2 type	1.70E-06
114	LI:1165056.1:2000SEP08	542	730	forward 2	KRAB	KRAB box	4.80E-31
114	LI:1165056.1:2000SEP08	1250	1318	forward 2	zf-C2H2	Zinc finger, C2H2 type	3.40E-08
115	LI:1175250.1:2000SEP08	157	309	forward 1	KRAB	KRAB box	1.10E-15
115	LI:1175250.1:2000SEP08	1141	1209	forward 1	zf-C2H2	Zinc finger, C2H2 type	5.70E-07
116	LI:1183192.1:2000SEP08	177	365	forward 3	KRAB	KRAB box	7.70E-35
116	LI:1183192.1:2000SEP08	1236	1304	forward 3	zf-C2H2	Zinc finger, C2H2 type	2.20E-06
117	LI:1183325.1:2000SEP08	259	444	forward 1	KRAB	KRAB box	2.80E-36



TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
117	LI:1183325.1:2000SEP08	1687	1755	forward 1	zf-C2H2	Zinc finger, C2H2 type	9.90E-07
118	LI:1178269.2:2000SEP08	150	332	forward 3	KRAB	KRAB box	2.80E-20
119	LI:813422.1:2000SEP08	118	288	forward 1	KRAB	KRAB box	1.80E-21
120	LI:1093049.6:2000SEP08	455	643	forward 2	KRAB	KRAB box	2.10E-38
122	LG:1041854.1:2000SEP08	489	629	forward 3	ATP-synt_DE	ATP synthase, Delta/Epsilon chain, long alpha-helix domain	1.50E-05
122	LG:1041854.1:2000SEP08	231	485	forward 3	ATP-synt_DE_N	ATP synthase, Delta/Epsilon chain, beta-sandwich domain	1.00E-29
126	LI:1078917.1:2000SEP08	8	421	forward 2	transferrin	Transferrin	1.60E-37
126	LI:1078917.1:2000SEP08	405	548	forward 3	transferrin	Transferrin	1.50E-11
127	LI:1012560.1:2000SEP08	372	884	forward 3	connexin	Connexin	6.50E-12
127	LI:1012560.1:2000SEP08	191	748	forward 2	connexin	Connexin	9.60E-11
131	LG:334345.1:2000SEP08	153	314	forward 3	trypsin	Trypsin	1.70E-21
131	LG:334345.1:2000SEP08	293	430	forward 2	trypsin	Trypsin	2.60E-21
133	LG:098580.1:2000SEP08	120	377	forward 3	ICE_p10	ICE-like protease (caspase) p10 domain	4.90E-31
134	LG:969572.1:2000SEP08	225	539	forward 3	cystatin	Cystatin domain	9.70E-38
135	LG:196958.1:2000SEP08	401	547	forward 2	trypsin	Trypsin	3.20E-23
135	LG:196958.1:2000SEP08	889	1029	forward 1	trypsin	Trypsin	1.60E-22
136	LG:1087811.1:2000SEP08	235	387	forward 1	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-trans isomerase	1.10E-26
136	LG:1087811.1:2000SEP08	396	539	forward 3	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-trans isomerase	5.50E-17
136	LG:1087811.1:2000SEP08	539	667	forward 2	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-trans isomerase	3.60E-10
137	LG:1327885.1:2000SEP08	93	554	forward 3	Peptidase_S26	Signal peptidase I	2.10E-61
138	LI:449393.1:2000SEP08	90	788	forward 3	cpn60_ICP1	TCP-1/cpn60 chaperonin family	9.80E-66
139	LI:897616.1:2000SEP08	183	467	forward 3	FKBP	FKBP-type peptidyl-prolyl cis-trans isomerases	1.10E-59
140	LI:736860.1:2000SEP08	58	237	forward 1	potato_inhibit	Potato inhibitor I family	3.70E-26
142	LI:1074263.1:2000SEP08	99	416	forward 3	cystatin	Cystatin domain	3.10E-43
143	LI:334345.1:2000SEP08	382	519	forward 1	trypsin	Trypsin	2.60E-21

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
143	LI:334345.1:2000SEP08	248	409	forward 2	trypsin	Trypsin	5.00E-20
145	LI:1188168.1:2000SEP08	309	1091	forward 3	Peptidase_C2	Calpain family cysteine protease	4.00E-10
146	LI:1065168.1:2000SEP08	226	408	forward 1	proteasome	Proteasome A-type and B-type	2.70E-16
146	LI:1065168.1:2000SEP08	113	226	forward 2	proteasome	Proteasome A-type and B-type	3.70E-08
147	LI:1180418.1:2000SEP08	235	387	forward 1	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-trans isomerase	1.10E-26
147	LI:1180418.1:2000SEP08	396	539	forward 3	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-trans isomerase	5.50E-17
147	LI:1180418.1:2000SEP08	539	667	forward 2	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-trans isomerase	3.60E-10
148	LG:232648.1:2000SEP08	277	774	forward 1	PseudolJ_synth_2	RNA pseudouridylylate synthase	4.20E-16
152	LG:241055.1:2000SEP08	742	972	forward 1	Transposase_1	Transposase	3.10E-26
154	LG:475629.1:2000SEP08	164	364	forward 2	Sm	Sm protein	5.20E-16
155	LI:348991.1:2000SEP08	106	801	forward 1	Fibrillarin	Fibrillarin	3.40E-167
156	LI:475629.1:2000SEP08	162	356	forward 3	Sm	Sm protein	7.70E-06
160	LI:758009.3:2000SEP08	202	375	forward 1	SH3	SH3 domain	1.70E-09
161	LG:331593.1:2000SEP08	306	476	forward 3	Ig	Immunoglobulin domain	6.80E-05
163	LI:814362.1:2000SEP08	138	380	forward 3	Ig	Immunoglobulin domain	2.70E-09
165	LI:726197.1:2000SEP08	60	242	forward 3	glutaredoxin	Glutaredoxin	8.30E-25
166	LI:1075314.1:2000SEP08	148	510	forward 1	Sdh_cyt	Succinate dehydrogenase cytochrome b subunit	6.90E-44
167	LI:437883.1:2000SEP08	176	640	forward 2	complex1_24kD	Respiratory-chain NADH dehydrogenase 24 Kd subunit	1.20E-14
167	LI:437883.1:2000SEP08	135	623	forward 3	complex1_24kD	Respiratory-chain NADH dehydrogenase 24 Kd subunit	3.60E-04
168	LG:336265.1:2000SEP08	60	239	forward 3	Collagen	Collagen triple helix repeat (20 copies)	2.20E-10
168	LG:336265.1:2000SEP08	439	612	forward 1	Collagen	Collagen triple helix repeat (20 copies)	3.50E-04
169	LG:407788.2:2000SEP08	984	1163	forward 3	Collagen	Collagen triple helix repeat (20 copies)	4.30E-07
171	LI:332655.2:2000SEP08	415	513	forward 1	ank	Ank repeat	3.10E-10
174	LG:362757.1:2000SEP08	97	480	forward 1	cofilin_ADF	Cofilin/tropomyosin-type actin-binding proteins	3.50E-47

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
175	LG:406770.1:2000SEP08	101	220	forward 2	GATA	GATA zinc finger	2.80E-20
177	LG:001929.1:2000SEP08	373	1314	forward 1	filament	Intermediate filament protein	1.60E-119
178	LI:401322.1:2000SEP08	161	271	forward 2	tubulin	Tubulin/FtsZ family	4.70E-14
180	LI:407242.1:2000SEP08	224	568	forward 2	BTB	BTB/POZ domain	3.40E-27
180	LI:407242.1:2000SEP08	1655	1798	forward 2	Kelch	Kelch motif	1.70E-10
181	LI:403409.1:2000SEP08	1458	1652	forward 3	FHA	FHA domain	3.00E-04
181	LI:403409.1:2000SEP08	78	1193	forward 3	kinesin	Kinesin motor domain	1.70E-170
182	LI:450798.1:2000SEP08	96	260	forward 3	tubulin	Tubulin/FtsZ family	2.40E-25
182	LI:450798.1:2000SEP08	266	484	forward 2	tubulin	Tubulin/FtsZ family	3.00E-20
183	LI:410317.1:2000SEP08	205	564	forward 1	filament	Intermediate filament protein	5.60E-32
183	LI:410317.1:2000SEP08	545	745	forward 2	filament	Intermediate filament protein	2.40E-29
183	LI:410317.1:2000SEP08	81	206	forward 3	filament	Intermediate filament protein	9.40E-08
184	LI:340268.1:2000SEP08	410	556	forward 2	actin	Actin	1.90E-19
184	LI:340268.1:2000SEP08	93	716	forward 3	actin	Actin	5.80E-17
184	LI:340268.1:2000SEP08	292	399	forward 1	actin	Actin	2.80E-13
187	LG:1043787.1:2000SEP08	109	432	forward 1	GSHPx	Glutathione peroxidase	4.50E-64
188	LG:1098931.16:2000SEP08	124	228	forward 1	Carboxyl_trans	Carboxyl transferase domain	3.00E-06
190	LI:1075297.1:2000SEP08	251	517	forward 2	MAPEG	MAPEG family	9.50E-37
192	LI:297070.1:2000SEP08	784	1083	forward 1	Branch	Core-2/1-Branching enzyme	3.60E-27
193	LI:1085041.1:2000SEP08	168	614	forward 3	lipocalin	Lipocalin / cytosolic fatty-acid binding protein family	1.20E-37
194	LI:1071544.1:2000SEP08	131	436	forward 2	NDK	Nucleoside diphosphate kinases	2.60E-76
194	LI:1071544.1:2000SEP08	436	567	forward 1	NDK	Nucleoside diphosphate kinases	9.30E-24
195	LI:2052480.1:2000SEP08	1000	1773	forward 1	hexokinase	Hexokinase	6.40E-99
195	LI:2052480.1:2000SEP08	513	1019	forward 3	hexokinase	Hexokinase	3.50E-51
195	LI:2052480.1:2000SEP08	155	514	forward 2	hexokinase	Hexokinase	1.40E-18
196	LG:450105.1:2000SEP08	86	490	forward 2	Ribosomal_S12	Ribosomal protein S12	6.60E-78
197	LG:450581.1:2000SEP08	82	276	forward 1	Ribosomal_S28e	Ribosomal protein S28e	5.00E-42
198	LG:450887.1:2000SEP08	48	344	forward 3	Ribosomal_L36e	Ribosomal protein L36e	6.90E-41
199	LG:460809.1:2000SEP08	3	236	forward 3	Ribosomal_L31e	Ribosomal protein L31e	6.00E-17
200	LG:452089.1:2000SEP08	107	268	forward 2	Ribosomal_L5	Ribosomal protein L5	2.30E-25

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
200	LG:452089.1:2000SEP08	278	577	forward 2	Ribosomal_L5_C	ribosomal L5P family C-terminus	2.60E-59
201	LG:1099416.1:2000SEP08	318	479	forward 3	Ribosomal_L37e	Ribosomal protein L37e	1.60E-13
202	LG:255713.1:2000SEP08	118	396	forward 1	Ribosomal_L31e	Ribosomal protein L31e	2.30E-14
203	LG:998903.1:2000SEP08	130	426	forward 1	Ribosomal_L7Ae	Ribosomal protein	8.40E-37
205	LG:1096907.1:2000SEP08	25	297	forward 1	Ribosomal_L37ae	L7Ae/L30e/S12e/Gadd45 family	3.20E-58
206	LG:1323741.1:2000SEP08	134	325	forward 2	Ribosomal_L29	Ribosomal L37ae protein family	1.70E-15
207	LG:1098372.1:2000SEP08	64	384	forward 1	Ribosomal_L22e	Ribosomal L29 protein	2.00E-27
207	LG:1098372.1:2000SEP08	53	391	forward 2	Ribosomal_L22e	Ribosomal L22e protein family	1.70E-06
208	LG:1006783.1:2000SEP08	16	312	forward 1	Ribosomal_L21e	Ribosomal L22e protein family	1.50E-69
209	LG:1097562.1:2000SEP08	137	430	forward 2	Ribosomal_L7Ae	Ribosomal protein L21e	2.40E-37
210	LG:998868.1:2000SEP08	32	667	forward 2	Ribosomal_S3Ae	Ribosomal protein	1.60E-145
212	LG:1400567.1:2000SEP08	38	271	forward 2	Ribosomal_L31e	Ribosomal S3Ae family	1.30E-07
213	LI:449404.1:2000SEP08	175	531	forward 1	Ribosomal_S11	Ribosomal protein L31e	6.90E-77
214	LI:449941.2:2000SEP08	61	438	forward 1	Ribosomal_S8e	Ribosomal protein S11	4.70E-84
215	LI:450229.1:2000SEP08	85	648	forward 1	Ribosomal_S7e	Ribosomal protein S8e	4.60E-83
216	LI:450399.3:2000SEP08	81	446	forward 3	Ribosomal_L14	Ribosomal protein S7e	1.20E-53
217	LI:455771.1:2000SEP08	69	473	forward 3	Ribosomal_S12	Ribosomal protein L14p/L23e	6.60E-78
218	LI:720459.1:2000SEP08	64	564	forward 1	Ribosomal_L6e	Ribosomal protein S12	8.80E-117
219	LI:723156.1:2000SEP08	96	380	forward 3	Ribosomal_L31e	Ribosomal protein L6e	8.40E-62
220	LI:728055.1:2000SEP08	81	479	forward 3	Ribosomal_S9	Ribosomal protein L31e	2.00E-89
221	LI:1020789.1:2000SEP08	70	441	forward 1	Ribosomal_S13	Ribosomal protein S9/S16	4.50E-73
222	LI:1071728.1:2000SEP08	139	426	forward 1	Ribosomal_S10	Ribosomal protein S13/S18	2.80E-54
223	LI:1084329.1:2000SEP08	200	439	forward 2	Ribosomal_L23	Ribosomal protein S10p/S20e	5.00E-32
224	LI:246422.1:2000SEP08	64	411	forward 1	Ribosomal_L22e	Ribosomal protein L23	3.20E-69
225	LI:1086066.1:2000SEP08	209	571	forward 2	Ribosomal_S7	Ribosomal L22e protein family	2.30E-25
227	LI:885368.1:2000SEP08	89	466	forward 2	Ribosomal_S8	Ribosomal protein S7p/S5e	1.60E-60
228	LI:481782.1:2000SEP08	263	550	forward 2	Ribosomal_S4e	Ribosomal protein S8	2.50E-27
229	LI:1093813.1:2000SEP08	43	234	forward 1	Ribosomal_L29	Ribosomal family S4e	1.70E-15
230	LI:449413.2:2000SEP08	90	494	forward 3	Ribosomal_S12	Ribosomal L29 protein	6.60E-78

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
231	LI:450105.1:2000SEP08	274	483	forward 1	Ribosomal_S12	Ribosomal protein S12	5.60E-37
231	LI:450105.1:2000SEP08	150	278	forward 3	Ribosomal_S12	Ribosomal protein S12	5.00E-15
231	LI:450105.1:2000SEP08	86	169	forward 2	Ribosomal_S12	Ribosomal protein S12	1.40E-06
234	LI:817330.1:2000SEP08	138	389	forward 3	Ribosomal_L7Ae	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family	1.50E-05
235	LI:817845.1:2000SEP08	83	502	forward 2	Ribosomal_S19e	Ribosomal protein S19e	2.10E-101
236	LI:460809.1:2000SEP08	1	306	forward 1	Ribosomal_L31e	Ribosomal protein L31e	4.20E-19
237	LI:815874.1:2000SEP08	237	464	forward 3	Ribosomal_S10	Ribosomal protein S10p/S20e	3.90E-27
239	LI:035973.1:2000SEP08	322	483	forward 1	Ribosomal_L37e	Ribosomal protein L37e	5.50E-09
241	LI:2049074.1:2000SEP08	225	569	forward 3	Ribosomal_L13e	Ribosomal protein L13e	2.80E-49
241	LI:2049074.1:2000SEP08	97	180	forward 1	Ribosomal_L13e	Ribosomal protein L13e	1.20E-12
241	LI:2049074.1:2000SEP08	44	229	forward 2	Ribosomal_L13e	Ribosomal protein L13e	1.10E-10
242	LI:1092460.1:2000SEP08	34	411	forward 1	histone	Core histone H2A/H2B/H3/H4	4.50E-47
244	LI:816655.2:2000SEP08	136	393	forward 1	HMG14_17	HMG14 and HMG17	5.50E-24
245	LG:414732.1:2000SEP08	79	534	forward 1	Glyco_hydro_2_N	Glycosyl hydrolases family 2, sugar binding domain	2.70E-10
247	LG:174022.1:2000SEP08	104	382	forward 2	ldh	lactate/malate dehydrogenase, NAD binding domain	1.10E-35
247	LG:174022.1:2000SEP08	363	446	forward 3	ldh	lactate/malate dehydrogenase, NAD binding domain	1.70E-08
247	LG:174022.1:2000SEP08	31	108	forward 1	ldh	lactate/malate dehydrogenase, NAD binding domain	2.40E-06
249	LI:414732.2:2000SEP08	79	534	forward 1	Glyco_hydro_2_N	Glycosyl hydrolases family 2, sugar binding domain	2.70E-10
255	LG:453922.1:2000SEP08	175	477	forward 1	SUI1	Translation Initiation factor SUI1	4.20E-58
256	LG:476342.3:2000SEP08	190	330	forward 1	Gamma-thlonin	Gamma-thlonins family	1.70E-19
257	LI:336801.1:2000SEP08	211	543	forward 1	SCP2	SCP-2 sterol transfer family	2.00E-28
259	LI:476342.1:2000SEP08	159	299	forward 3	Gamma-thlonin	Gamma-thlonins family	1.70E-19
260	LI:1072804.1:2000SEP08	278	481	forward 2	pp-binding	Phosphopantetheine attachment site	3.90E-14
261	LI:455450.1:2000SEP08	1	426	forward 1	dehydrin	Dehydrin	4.20E-41
262	LI:1073699.1:2000SEP08	51	248	forward 3	metallo	Metallothionein	1.70E-25

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
263	LI:1013729.1:2000SEP08	33	392	forward 3	lipocalin	Lipocalin / cytosolic fatty-acid binding protein family	1.30E-24
264	LI:2050322.2:2000SEP08	109	318	forward 1	rrm	RNA recognition motiff. (a.k.a. RRM, RBD, or RNP domain)	5.00E-24
265	LI:891327.1:2000SEP08	303	551	forward 3	HIN	HIN-200/IF120x domain	1.60E-11
267	LG:220085.1:2000SEP08	228	521	forward 3	rvp	Retroviral aspartyl protease	2.10E-15
267	LG:220085.1:2000SEP08	728	853	forward 2	rvf	Reverse transcriptase (RNA-dependent DNA polymerase)	1.10E-04
268	LG:406709.1:2000SEP08	198	611	forward 3	GTP_CDC	Cell division protein	3.70E-70
268	LG:406709.1:2000SEP08	635	1018	forward 2	GTP_CDC	Cell division protein	1.30E-44
270	LI:1073027.1:2000SEP08	16	525	forward 1	ThlJ	ThlJ/PfpI family	6.10E-54
271	LI:347635.1:2000SEP08	515	610	forward 2	EGF	EGF-like domain	6.70E-09
271	LI:347635.1:2000SEP08	790	894	forward 1	EGF	EGF-like domain	1.10E-04
271	LI:347635.1:2000SEP08	1251	1346	forward 3	EGF	EGF-like domain	1.20E-04
273	LI:406709.1:2000SEP08	374	835	forward 2	GTP_CDC	Cell division protein	5.20E-39
273	LI:406709.1:2000SEP08	852	1031	forward 3	GTP_CDC	Cell division protein	7.60E-27
275	LI:213208.1:2000SEP08	279	491	forward 3	K_tetra	K+ channel tetramerisation domain	1.10E-07

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
1	LG:405741.3:2000SEP08	76	162	forward 1	TM	N in
1	LG:405741.3:2000SEP08	499	585	forward 1	TM	N in
1	LG:405741.3:2000SEP08	1000	1086	forward 1	TM	N in
1	LG:405741.3:2000SEP08	647	733	forward 2	TM	N out
1	LG:405741.3:2000SEP08	995	1081	forward 2	TM	N out
1	LG:405741.3:2000SEP08	537	593	forward 3	TM	N in
1	LG:405741.3:2000SEP08	1029	1115	forward 3	TM	N in
2	LG:337194.1:2000SEP08	208	294	forward 1	TM	
2	LG:337194.1:2000SEP08	331	399	forward 1	TM	
2	LG:337194.1:2000SEP08	1129	1215	forward 1	TM	
2	LG:337194.1:2000SEP08	1501	1551	forward 1	TM	
2	LG:337194.1:2000SEP08	1639	1710	forward 1	TM	
2	LG:337194.1:2000SEP08	1801	1887	forward 1	TM	
2	LG:337194.1:2000SEP08	1948	1998	forward 1	TM	
2	LG:337194.1:2000SEP08	188	271	forward 2	TM	N in
2	LG:337194.1:2000SEP08	1472	1540	forward 2	TM	N in
2	LG:337194.1:2000SEP08	1568	1624	forward 2	TM	N in
2	LG:337194.1:2000SEP08	1622	1708	forward 2	TM	N in
2	LG:337194.1:2000SEP08	1793	1855	forward 2	TM	N in
2	LG:337194.1:2000SEP08	1895	1957	forward 2	TM	N in
2	LG:337194.1:2000SEP08	1988	2050	forward 2	TM	N in
2	LG:337194.1:2000SEP08	2081	2143	forward 2	TM	N in
2	LG:337194.1:2000SEP08	204	290	forward 3	TM	N in
2	LG:337194.1:2000SEP08	339	392	forward 3	TM	N in
2	LG:337194.1:2000SEP08	678	764	forward 3	TM	N in
2	LG:337194.1:2000SEP08	813	899	forward 3	TM	N in
2	LG:337194.1:2000SEP08	1086	1172	forward 3	TM	N in
2	LG:337194.1:2000SEP08	1455	1541	forward 3	TM	N in
2	LG:337194.1:2000SEP08	1848	1913	forward 3	TM	N in
2	LG:337194.1:2000SEP08	2106	2159	forward 3	TM	N in
4	LG:372569.5:2000SEP08	400	465	forward 1	TM	
4	LG:372569.5:2000SEP08	544	594	forward 1	TM	
4	LG:372569.5:2000SEP08	985	1044	forward 1	TM	
4	LG:372569.5:2000SEP08	135	221	forward 3	TM	N out
5	LG:968765.1:2000SEP08	131	190	forward 2	TM	N out
6	LG:255999.16:2000SEP08	221	277	forward 2	TM	N out
7	LG:977820.9:2000SEP08	1267	1338	forward 1	TM	N out
8	LI:1071608.1:2000SEP08	31	117	forward 1	TM	N in
8	LI:1071608.1:2000SEP08	319	405	forward 1	TM	N in
8	LI:1071608.1:2000SEP08	108	155	forward 3	TM	N out
10	LI:453570.1:2000SEP08	361	447	forward 1	TM	
11	LI:072072.1:2000SEP08	325	396	forward 1	TM	N out
11	LI:072072.1:2000SEP08	664	750	forward 1	TM	N out
11	LI:072072.1:2000SEP08	1489	1566	forward 1	TM	N out
11	LI:072072.1:2000SEP08	1624	1710	forward 1	TM	N out
11	LI:072072.1:2000SEP08	1810	1872	forward 1	TM	N out
11	LI:072072.1:2000SEP08	1900	1962	forward 1	TM	N out
11	LI:072072.1:2000SEP08	2119	2187	forward 1	TM	N out
11	LI:072072.1:2000SEP08	227	313	forward 2	TM	

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
11	LI:072072.1:2000SEP08	620	682	forward 2	TM	
11	LI:072072.1:2000SEP08	695	757	forward 2	TM	
11	LI:072072.1:2000SEP08	800	886	forward 2	TM	
11	LI:072072.1:2000SEP08	1508	1594	forward 2	TM	
11	LI:072072.1:2000SEP08	1658	1744	forward 2	TM	
11	LI:072072.1:2000SEP08	1931	2017	forward 2	TM	
11	LI:072072.1:2000SEP08	2699	2782	forward 2	TM	
11	LI:072072.1:2000SEP08	312	398	forward 3	TM	N out
11	LI:072072.1:2000SEP08	969	1055	forward 3	TM	N out
11	LI:072072.1:2000SEP08	1449	1508	forward 3	TM	N out
11	LI:072072.1:2000SEP08	1674	1745	forward 3	TM	N out
11	LI:072072.1:2000SEP08	1917	1985	forward 3	TM	N out
12	LI:148565.4:2000SEP08	1350	1436	forward 3	TM	N out
13	LI:368626.4:2000SEP08	1219	1287	forward 1	TM	N out
13	LI:368626.4:2000SEP08	590	673	forward 2	TM	N out
13	LI:368626.4:2000SEP08	671	739	forward 2	TM	N out
13	LI:368626.4:2000SEP08	908	970	forward 2	TM	N out
13	LI:368626.4:2000SEP08	108	167	forward 3	TM	N out
13	LI:368626.4:2000SEP08	777	863	forward 3	TM	N out
14	LI:346123.1:2000SEP08	199	246	forward 1	TM	N in
14	LI:346123.1:2000SEP08	738	824	forward 3	TM	N out
14	LI:346123.1:2000SEP08	981	1049	forward 3	TM	N out
15	LI:335795.11:2000SEP08	1483	1551	forward 1	TM	N in
15	LI:335795.11:2000SEP08	2791	2841	forward 1	TM	N in
15	LI:335795.11:2000SEP08	2965	3030	forward 1	TM	N in
15	LI:335795.11:2000SEP08	1691	1747	forward 2	TM	N in
15	LI:335795.11:2000SEP08	1952	2038	forward 2	TM	N in
15	LI:335795.11:2000SEP08	2492	2566	forward 2	TM	N in
15	LI:335795.11:2000SEP08	3035	3097	forward 2	TM	N in
15	LI:335795.11:2000SEP08	1665	1751	forward 3	TM	N in
15	LI:335795.11:2000SEP08	2418	2504	forward 3	TM	N in
15	LI:335795.11:2000SEP08	2742	2822	forward 3	TM	N in
16	LI:246023.2:2000SEP08	244	303	forward 1	TM	N out
16	LI:246023.2:2000SEP08	439	489	forward 1	TM	N out
17	LG:1100661.1:2000SEP08	21	83	forward 3	TM	N out
18	LG:475856.1:2000SEP08	454	540	forward 1	TM	N out
18	LG:475856.1:2000SEP08	29	106	forward 2	TM	N in
18	LG:475856.1:2000SEP08	449	535	forward 2	TM	N in
18	LG:475856.1:2000SEP08	447	533	forward 3	TM	
19	LG:1015343.1:2000SEP08	97	159	forward 1	TM	
19	LG:1015343.1:2000SEP08	431	499	forward 2	TM	N in
19	LG:1015343.1:2000SEP08	411	482	forward 3	TM	N out
22	LG:213947.1:2000SEP08	100	186	forward 1	TM	N out
22	LG:213947.1:2000SEP08	244	330	forward 1	TM	N out
23	LI:720641.1:2000SEP08	664	732	forward 1	TM	N out
23	LI:720641.1:2000SEP08	197	262	forward 2	TM	N in
23	LI:720641.1:2000SEP08	293	352	forward 2	TM	N in
23	LI:720641.1:2000SEP08	413	484	forward 2	TM	N in
23	LI:720641.1:2000SEP08	533	619	forward 2	TM	N in



TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
23	LI:720641.1:2000SEP08	683	769	forward 2	TM	N in
23	LI:720641.1:2000SEP08	920	979	forward 2	TM	N in
23	LI:720641.1:2000SEP08	132	218	forward 3	TM	N out
23	LI:720641.1:2000SEP08	519	569	forward 3	TM	N out
23	LI:720641.1:2000SEP08	642	728	forward 3	TM	N out
25	LI:734904.1:2000SEP08	58	132	forward 1	TM	N out
25	LI:734904.1:2000SEP08	196	243	forward 1	TM	N out
25	LI:734904.1:2000SEP08	508	573	forward 1	TM	N out
25	LI:734904.1:2000SEP08	886	972	forward 1	TM	N out
25	LI:734904.1:2000SEP08	50	136	forward 2	TM	N out
25	LI:734904.1:2000SEP08	578	655	forward 2	TM	N out
25	LI:734904.1:2000SEP08	1277	1342	forward 2	TM	N out
25	LI:734904.1:2000SEP08	1491	1547	forward 3	TM	N out
26	LI:1178118.1:2000SEP08	1054	1140	forward 1	TM	N in
26	LI:1178118.1:2000SEP08	1411	1488	forward 1	TM	N in
26	LI:1178118.1:2000SEP08	1612	1662	forward 1	TM	N in
26	LI:1178118.1:2000SEP08	1891	1977	forward 1	TM	N in
26	LI:1178118.1:2000SEP08	2041	2127	forward 1	TM	N in
26	LI:1178118.1:2000SEP08	2398	2475	forward 1	TM	N in
26	LI:1178118.1:2000SEP08	1394	1468	forward 2	TM	N in
26	LI:1178118.1:2000SEP08	1514	1600	forward 2	TM	N in
26	LI:1178118.1:2000SEP08	1883	1945	forward 2	TM	N in
26	LI:1178118.1:2000SEP08	2351	2437	forward 2	TM	N in
26	LI:1178118.1:2000SEP08	1641	1703	forward 3	TM	N in
26	LI:1178118.1:2000SEP08	1995	2081	forward 3	TM	N in
26	LI:1178118.1:2000SEP08	2472	2534	forward 3	TM	N in
27	LI:213947.1:2000SEP08	63	149	forward 3	TM	N out
27	LI:213947.1:2000SEP08	198	284	forward 3	TM	N out
28	LG:407304.1:2000SEP08	157	243	forward 1	TM	N out
28	LG:407304.1:2000SEP08	107	193	forward 2	TM	N in
29	LG:337358.1:2000SEP08	2452	2508	forward 1	TM	N in
29	LG:337358.1:2000SEP08	2767	2853	forward 1	TM	N in
29	LG:337358.1:2000SEP08	884	970	forward 2	TM	
29	LG:337358.1:2000SEP08	1556	1642	forward 2	TM	
29	LG:337358.1:2000SEP08	2096	2173	forward 2	TM	
29	LG:337358.1:2000SEP08	2756	2842	forward 2	TM	
29	LG:337358.1:2000SEP08	2730	2816	forward 3	TM	N in
30	LG:986090.1:2000SEP08	583	630	forward 1	TM	N in
30	LG:986090.1:2000SEP08	197	271	forward 2	TM	N out
30	LG:986090.1:2000SEP08	521	607	forward 2	TM	N out
31	LG:123250.1:2000SEP08	1223	1309	forward 2	TM	N out
32	LG:1028774.2:2000SEP08	313	399	forward 1	TM	N in
32	LG:1028774.2:2000SEP08	934	1020	forward 1	TM	N in
32	LG:1028774.2:2000SEP08	554	640	forward 2	TM	N out
32	LG:1028774.2:2000SEP08	1010	1093	forward 2	TM	N out
32	LG:1028774.2:2000SEP08	1154	1240	forward 2	TM	N out
33	LG:338927.6:2000SEP08	17	88	forward 2	TM	N out
34	LG:332944.2:2000SEP08	13	66	forward 1	TM	N in
34	LG:332944.2:2000SEP08	1609	1659	forward 1	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
34	LG:332944.2:2000SEP08	1852	1938	forward 1	TM	N in
34	LG:332944.2:2000SEP08	1154	1240	forward 2	TM	N out
34	LG:332944.2:2000SEP08	2339	2407	forward 2	TM	N out
34	LG:332944.2:2000SEP08	1785	1856	forward 3	TM	N in
35	LI:347174.5:2000SEP08	640	702	forward 1	TM	N in
35	LI:347174.5:2000SEP08	721	783	forward 1	TM	N in
35	LI:347174.5:2000SEP08	1330	1407	forward 1	TM	N in
35	LI:347174.5:2000SEP08	743	808	forward 2	TM	N in
35	LI:347174.5:2000SEP08	854	940	forward 2	TM	N in
35	LI:347174.5:2000SEP08	753	839	forward 3	TM	
35	LI:347174.5:2000SEP08	873	959	forward 3	TM	
37	LI:723144.1:2000SEP08	396	473	forward 3	TM	N out
40	LI:284797.3:2000SEP08	994	1044	forward 1	TM	N in
40	LI:284797.3:2000SEP08	1426	1476	forward 1	TM	N in
40	LI:284797.3:2000SEP08	1564	1632	forward 1	TM	N in
40	LI:284797.3:2000SEP08	998	1060	forward 2	TM	N in
40	LI:284797.3:2000SEP08	1796	1858	forward 2	TM	N in
40	LI:284797.3:2000SEP08	735	821	forward 3	TM	N in
40	LI:284797.3:2000SEP08	969	1028	forward 3	TM	N in
40	LI:284797.3:2000SEP08	1422	1508	forward 3	TM	N in
41	LI:1092901.1:2000SEP08	92	160	forward 2	TM	N out
41	LI:1092901.1:2000SEP08	473	559	forward 2	TM	N out
41	LI:1092901.1:2000SEP08	495	578	forward 3	TM	N in
42	LI:228930.1:2000SEP08	141	197	forward 3	TM	N in
43	LI:722913.1:2000SEP08	142	201	forward 1	TM	N in
43	LI:722913.1:2000SEP08	485	541	forward 2	TM	N out
44	LG:457478.1:2000SEP08	454	540	forward 1	TM	N out
44	LG:457478.1:2000SEP08	282	329	forward 3	TM	
45	LG:358719.1:2000SEP08	109	195	forward 1	TM	N in
45	LG:358719.1:2000SEP08	346	426	forward 1	TM	N in
45	LG:358719.1:2000SEP08	691	762	forward 1	TM	N in
45	LG:358719.1:2000SEP08	772	852	forward 1	TM	N in
45	LG:358719.1:2000SEP08	913	969	forward 1	TM	N in
45	LG:358719.1:2000SEP08	35	97	forward 2	TM	
45	LG:358719.1:2000SEP08	125	187	forward 2	TM	
46	LG:105160.5:2000SEP08	205	255	forward 1	TM	N out
46	LG:105160.5:2000SEP08	737	793	forward 2	TM	N in
47	LG:400705.1:2000SEP08	871	948	forward 1	TM	N out
47	LG:400705.1:2000SEP08	867	953	forward 3	TM	N out
48	LG:221977.1:2000SEP08	76	126	forward 1	TM	N out
48	LG:221977.1:2000SEP08	1208	1276	forward 2	TM	N out
49	LG:898771.1:2000SEP08	910	996	forward 1	TM	N out
49	LG:898771.1:2000SEP08	767	823	forward 2	TM	N out
49	LG:898771.1:2000SEP08	905	973	forward 2	TM	N out
49	LG:898771.1:2000SEP08	894	962	forward 3	TM	N in
50	LI:457478.1:2000SEP08	284	331	forward 2	TM	N out
50	LI:457478.1:2000SEP08	456	542	forward 3	TM	N out
53	LI:888730.1:2000SEP08	55	123	forward 1	TM	N out
53	LI:888730.1:2000SEP08	262	324	forward 1	TM	N out

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
54	LI:358719.1:2000SEP08	742	804	forward 1	TM	N out
54	LI:358719.1:2000SEP08	820	882	forward 1	TM	N out
54	LI:358719.1:2000SEP08	20	106	forward 2	TM	N in
54	LI:358719.1:2000SEP08	152	205	forward 2	TM	N in
54	LI:358719.1:2000SEP08	953	1009	forward 2	TM	N in
54	LI:358719.1:2000SEP08	132	197	forward 3	TM	N out
55	LI:351342.3:2000SEP08	10	63	forward 1	TM	N out
55	LI:351342.3:2000SEP08	142	228	forward 1	TM	N out
56	LI:256099.2:2000SEP08	1609	1683	forward 1	TM	N out
57	LI:2051991.1:2000SEP08	788	844	forward 2	TM	N out
57	LI:2051991.1:2000SEP08	153	215	forward 3	TM	N out
57	LI:2051991.1:2000SEP08	225	287	forward 3	TM	N out
57	LI:2051991.1:2000SEP08	1077	1157	forward 3	TM	N out
58	LG:980769.1:2000SEP08	47	130	forward 2	TM	N out
58	LG:980769.1:2000SEP08	260	346	forward 2	TM	N out
58	LG:980769.1:2000SEP08	99	161	forward 3	TM	N out
58	LG:980769.1:2000SEP08	177	239	forward 3	TM	N out
58	LG:980769.1:2000SEP08	714	800	forward 3	TM	N out
59	LG:332474.3:2000SEP08	308	361	forward 2	TM	
59	LG:332474.3:2000SEP08	956	1024	forward 2	TM	
59	LG:332474.3:2000SEP08	963	1019	forward 3	TM	N out
60	LG:1087707.1:2000SEP08	862	933	forward 1	TM	
60	LG:1087707.1:2000SEP08	994	1062	forward 1	TM	
60	LG:1087707.1:2000SEP08	1066	1134	forward 1	TM	
60	LG:1087707.1:2000SEP08	620	706	forward 2	TM	N out
60	LG:1087707.1:2000SEP08	749	814	forward 2	TM	N out
60	LG:1087707.1:2000SEP08	974	1036	forward 2	TM	N out
60	LG:1087707.1:2000SEP08	1052	1114	forward 2	TM	N out
60	LG:1087707.1:2000SEP08	696	782	forward 3	TM	N in
60	LG:1087707.1:2000SEP08	972	1034	forward 3	TM	N in
60	LG:1087707.1:2000SEP08	1047	1109	forward 3	TM	N in
61	LG:415349.1:2000SEP08	310	396	forward 1	TM	N out
61	LG:415349.1:2000SEP08	463	534	forward 1	TM	N out
61	LG:415349.1:2000SEP08	296	358	forward 2	TM	N out
61	LG:415349.1:2000SEP08	374	436	forward 2	TM	N out
61	LG:415349.1:2000SEP08	90	176	forward 3	TM	N out
61	LG:415349.1:2000SEP08	213	263	forward 3	TM	N out
61	LG:415349.1:2000SEP08	273	323	forward 3	TM	N out
61	LG:415349.1:2000SEP08	336	392	forward 3	TM	N out
62	LG:132420.2:2000SEP08	1162	1248	forward 1	TM	N in
62	LG:132420.2:2000SEP08	1359	1424	forward 3	TM	N in
63	LG:394201.1:2000SEP08	10	96	forward 1	TM	N in
63	LG:394201.1:2000SEP08	520	606	forward 1	TM	N in
63	LG:394201.1:2000SEP08	23	82	forward 2	TM	N out
63	LG:394201.1:2000SEP08	161	226	forward 2	TM	N out
63	LG:394201.1:2000SEP08	326	388	forward 2	TM	N out
63	LG:394201.1:2000SEP08	404	466	forward 2	TM	N out
63	LG:394201.1:2000SEP08	495	569	forward 3	TM	N out
64	LG:1060884.1:2000SEP08	19	105	forward 1	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
64	LG:1060884.1:2000SEP08	595	681	forward 1	TM	N In
64	LG:1060884.1:2000SEP08	17	103	forward 2	TM	N In
64	LG:1060884.1:2000SEP08	101	166	forward 2	TM	N In
64	LG:1060884.1:2000SEP08	236	301	forward 2	TM	N In
64	LG:1060884.1:2000SEP08	42	110	forward 3	TM	N out
68	LG:979390.2:2000SEP08	190	240	forward 1	TM	N out
68	LG:979390.2:2000SEP08	976	1029	forward 1	TM	N out
68	LG:979390.2:2000SEP08	857	919	forward 2	TM	
69	LG:1400447.1:2000SEP08	209	280	forward 2	TM	N In
70	LG:1400562.1:2000SEP08	337	411	forward 1	TM	N out
70	LG:1400562.1:2000SEP08	433	519	forward 1	TM	N out
70	LG:1400562.1:2000SEP08	573	629	forward 3	TM	N In
71	LG:1076130.1:2000SEP08	481	531	forward 1	TM	N out
72	LG:1064459.1:2000SEP08	747	833	forward 3	TM	N out
73	LG:1079415.14:2000SEP08	86	172	forward 2	TM	N out
73	LG:1079415.14:2000SEP08	24	110	forward 3	TM	N out
75	LG:1088431.2:2000SEP08	403	465	forward 1	TM	N out
75	LG:1088431.2:2000SEP08	378	455	forward 3	TM	N out
76	LG:1329462.2:2000SEP08	55	141	forward 1	TM	N out
76	LG:1329462.2:2000SEP08	313	366	forward 1	TM	N out
76	LG:1329462.2:2000SEP08	300	383	forward 3	TM	N In
78	LI:722577.1:2000SEP08	236	310	forward 2	TM	N out
78	LI:722577.1:2000SEP08	320	406	forward 2	TM	N out
78	LI:722577.1:2000SEP08	9	95	forward 3	TM	N In
78	LI:722577.1:2000SEP08	333	407	forward 3	TM	N In
79	LI:322783.16:2000SEP08	169	231	forward 1	TM	N In
79	LI:322783.16:2000SEP08	256	318	forward 1	TM	N In
79	LI:322783.16:2000SEP08	218	304	forward 2	TM	N out
79	LI:322783.16:2000SEP08	326	412	forward 2	TM	N out
79	LI:322783.16:2000SEP08	72	134	forward 3	TM	N In
79	LI:322783.16:2000SEP08	150	212	forward 3	TM	N In
79	LI:322783.16:2000SEP08	270	356	forward 3	TM	N In
81	LI:038859.2:2000SEP08	679	765	forward 1	TM	N out
81	LI:038859.2:2000SEP08	47	130	forward 2	TM	N out
81	LI:038859.2:2000SEP08	260	346	forward 2	TM	N out
81	LI:038859.2:2000SEP08	99	161	forward 3	TM	N out
81	LI:038859.2:2000SEP08	177	239	forward 3	TM	N out
82	LI:1046117.1:2000SEP08	351	431	forward 3	TM	N out
85	LI:1170585.2:2000SEP08	141	203	forward 3	TM	N out
86	LI:719531.2:2000SEP08	81	134	forward 3	TM	N out
90	LI:1091881.1:2000SEP08	53	133	forward 2	TM	N out
93	LI:1166249.1:2000SEP08	11	88	forward 2	TM	N In
93	LI:1166249.1:2000SEP08	395	481	forward 2	TM	N In
94	LI:799675.1:2000SEP08	25	99	forward 1	TM	N out
94	LI:799675.1:2000SEP08	403	480	forward 1	TM	N out
94	LI:799675.1:2000SEP08	98	184	forward 2	TM	N out
94	LI:799675.1:2000SEP08	383	442	forward 2	TM	N out
94	LI:799675.1:2000SEP08	120	203	forward 3	TM	N out
95	LI:1178899.1:2000SEP08	73	159	forward 1	TM	N out

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
95	LI:1178899.1:2000SEP08	331	384	forward 1	TM	N out
95	LI:1178899.1:2000SEP08	315	398	forward 3	TM	N out
100	LI:805158.1:2000SEP08	22	99	forward 1	TM	
102	LI:1174107.2:2000SEP08	300	362	forward 3	TM	N out
102	LI:1174107.2:2000SEP08	375	437	forward 3	TM	N out
103	LI:1177434.2:2000SEP08	1023	1091	forward 3	TM	N out
104	LI:1184255.1:2000SEP08	611	694	forward 2	TM	N out
107	LI:1166752.1:2000SEP08	31	93	forward 1	TM	N in
107	LI:1166752.1:2000SEP08	112	174	forward 1	TM	N in
107	LI:1166752.1:2000SEP08	193	255	forward 1	TM	N in
107	LI:1166752.1:2000SEP08	263	340	forward 2	TM	N out
107	LI:1166752.1:2000SEP08	33	110	forward 3	TM	N in
107	LI:1166752.1:2000SEP08	117	173	forward 3	TM	N in
107	LI:1166752.1:2000SEP08	219	305	forward 3	TM	N in
109	LI:242665.2:2000SEP08	81	137	forward 3	TM	N out
110	LI:208637.1:2000SEP08	304	390	forward 1	TM	N out
110	LI:208637.1:2000SEP08	2449	2535	forward 1	TM	N out
110	LI:208637.1:2000SEP08	2983	3069	forward 1	TM	N out
110	LI:208637.1:2000SEP08	3439	3525	forward 1	TM	N out
110	LI:208637.1:2000SEP08	3538	3624	forward 1	TM	N out
110	LI:208637.1:2000SEP08	3793	3879	forward 1	TM	N out
110	LI:208637.1:2000SEP08	3889	3945	forward 1	TM	N out
110	LI:208637.1:2000SEP08	4081	4152	forward 1	TM	N out
110	LI:208637.1:2000SEP08	4159	4212	forward 1	TM	N out
110	LI:208637.1:2000SEP08	4435	4521	forward 1	TM	N out
110	LI:208637.1:2000SEP08	4900	4953	forward 1	TM	N out
110	LI:208637.1:2000SEP08	278	340	forward 2	TM	N out
110	LI:208637.1:2000SEP08	353	415	forward 2	TM	N out
110	LI:208637.1:2000SEP08	503	553	forward 2	TM	N out
110	LI:208637.1:2000SEP08	566	628	forward 2	TM	N out
110	LI:208637.1:2000SEP08	653	715	forward 2	TM	N out
110	LI:208637.1:2000SEP08	764	850	forward 2	TM	N out
110	LI:208637.1:2000SEP08	899	952	forward 2	TM	N out
110	LI:208637.1:2000SEP08	1721	1804	forward 2	TM	N out
110	LI:208637.1:2000SEP08	2393	2443	forward 2	TM	N out
110	LI:208637.1:2000SEP08	2903	2977	forward 2	TM	N out
110	LI:208637.1:2000SEP08	3011	3082	forward 2	TM	N out
110	LI:208637.1:2000SEP08	3119	3205	forward 2	TM	N out
110	LI:208637.1:2000SEP08	3299	3370	forward 2	TM	N out
110	LI:208637.1:2000SEP08	3404	3478	forward 2	TM	N out
110	LI:208637.1:2000SEP08	4238	4291	forward 2	TM	N out
110	LI:208637.1:2000SEP08	4301	4375	forward 2	TM	N out
110	LI:208637.1:2000SEP08	381	431	forward 3	TM	N in
110	LI:208637.1:2000SEP08	2094	2165	forward 3	TM	N in
110	LI:208637.1:2000SEP08	2547	2630	forward 3	TM	N in
110	LI:208637.1:2000SEP08	2877	2948	forward 3	TM	N in
110	LI:208637.1:2000SEP08	2982	3068	forward 3	TM	N in
110	LI:208637.1:2000SEP08	3186	3272	forward 3	TM	N in
110	LI:208637.1:2000SEP08	3273	3359	forward 3	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
110	LI:208637.1:2000SEP08	3369	3431	forward 3	TM	N in
110	LI:208637.1:2000SEP08	3453	3515	forward 3	TM	N in
110	LI:208637.1:2000SEP08	3528	3599	forward 3	TM	N in
110	LI:208637.1:2000SEP08	3666	3752	forward 3	TM	N in
110	LI:208637.1:2000SEP08	3930	4016	forward 3	TM	N in
110	LI:208637.1:2000SEP08	4242	4328	forward 3	TM	N in
110	LI:208637.1:2000SEP08	4440	4526	forward 3	TM	N in
110	LI:208637.1:2000SEP08	4695	4775	forward 3	TM	N in
111	LI:2051808.1:2000SEP08	775	852	forward 1	TM	N in
111	LI:2051808.1:2000SEP08	758	844	forward 2	TM	N in
113	LI:1177337.1:2000SEP08	1639	1713	forward 1	TM	N out
113	LI:1177337.1:2000SEP08	846	932	forward 3	TM	N out
113	LI:1177337.1:2000SEP08	1620	1670	forward 3	TM	N out
114	LI:1165056.1:2000SEP08	918	1004	forward 3	TM	N out
115	LI:1175250.1:2000SEP08	596	652	forward 2	TM	N in
115	LI:1175250.1:2000SEP08	917	1003	forward 2	TM	N in
116	LI:1183192.1:2000SEP08	1405	1491	forward 1	TM	N out
116	LI:1183192.1:2000SEP08	1319	1378	forward 2	TM	N in
116	LI:1183192.1:2000SEP08	1449	1529	forward 3	TM	
117	LI:1183325.1:2000SEP08	554	607	forward 2	TM	N in
119	LI:813422.1:2000SEP08	61	138	forward 1	TM	N out
119	LI:813422.1:2000SEP08	493	573	forward 1	TM	N out
119	LI:813422.1:2000SEP08	682	753	forward 1	TM	N out
119	LI:813422.1:2000SEP08	1123	1185	forward 1	TM	N out
119	LI:813422.1:2000SEP08	1198	1260	forward 1	TM	N out
119	LI:813422.1:2000SEP08	1555	1632	forward 1	TM	N out
119	LI:813422.1:2000SEP08	692	751	forward 2	TM	N out
119	LI:813422.1:2000SEP08	1106	1183	forward 2	TM	N out
119	LI:813422.1:2000SEP08	1553	1615	forward 2	TM	N out
119	LI:813422.1:2000SEP08	57	143	forward 3	TM	
119	LI:813422.1:2000SEP08	255	341	forward 3	TM	
119	LI:813422.1:2000SEP08	492	575	forward 3	TM	
119	LI:813422.1:2000SEP08	585	635	forward 3	TM	
119	LI:813422.1:2000SEP08	705	779	forward 3	TM	
119	LI:813422.1:2000SEP08	1113	1175	forward 3	TM	
119	LI:813422.1:2000SEP08	1452	1499	forward 3	TM	
120	LI:1093049.6:2000SEP08	25	87	forward 1	TM	N in
120	LI:1093049.6:2000SEP08	121	183	forward 1	TM	N in
120	LI:1093049.6:2000SEP08	424	480	forward 1	TM	N in
120	LI:1093049.6:2000SEP08	925	1011	forward 1	TM	N in
120	LI:1093049.6:2000SEP08	1069	1152	forward 1	TM	N in
120	LI:1093049.6:2000SEP08	1204	1290	forward 1	TM	N in
120	LI:1093049.6:2000SEP08	92	142	forward 2	TM	N in
120	LI:1093049.6:2000SEP08	815	883	forward 2	TM	N in
120	LI:1093049.6:2000SEP08	968	1039	forward 2	TM	N in
120	LI:1093049.6:2000SEP08	1235	1318	forward 2	TM	N in
120	LI:1093049.6:2000SEP08	27	89	forward 3	TM	N in
120	LI:1093049.6:2000SEP08	105	167	forward 3	TM	N in
120	LI:1093049.6:2000SEP08	939	1010	forward 3	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
120	LI:1093049.6:2000SEP08	1023	1109	forward 3	TM	N in
120	LI:1093049.6:2000SEP08	1209	1286	forward 3	TM	N in
121	LI:202192.4:2000SEP08	199	273	forward 1	TM	N out
121	LI:202192.4:2000SEP08	457	531	forward 1	TM	N out
122	LG:1041854.1:2000SEP08	285	371	forward 3	TM	N in
123	LG:1100502.1:2000SEP08	314	400	forward 2	TM	N in
123	LG:1100502.1:2000SEP08	270	338	forward 3	TM	N in
124	LI:726414.1:2000SEP08	172	243	forward 1	TM	N out
124	LI:726414.1:2000SEP08	551	637	forward 2	TM	N out
125	LI:400517.4:2000SEP08	337	399	forward 1	TM	N out
125	LI:400517.4:2000SEP08	409	471	forward 1	TM	N out
125	LI:400517.4:2000SEP08	775	855	forward 1	TM	N out
125	LI:400517.4:2000SEP08	994	1044	forward 1	TM	N out
125	LI:400517.4:2000SEP08	1105	1191	forward 1	TM	N out
125	LI:400517.4:2000SEP08	1825	1911	forward 1	TM	N out
125	LI:400517.4:2000SEP08	2239	2325	forward 1	TM	N out
125	LI:400517.4:2000SEP08	2860	2913	forward 1	TM	N out
125	LI:400517.4:2000SEP08	89	151	forward 2	TM	N out
125	LI:400517.4:2000SEP08	182	244	forward 2	TM	N out
125	LI:400517.4:2000SEP08	311	388	forward 2	TM	N out
125	LI:400517.4:2000SEP08	479	565	forward 2	TM	N out
125	LI:400517.4:2000SEP08	653	739	forward 2	TM	N out
125	LI:400517.4:2000SEP08	791	874	forward 2	TM	N out
125	LI:400517.4:2000SEP08	902	988	forward 2	TM	N out
125	LI:400517.4:2000SEP08	2249	2329	forward 2	TM	N out
125	LI:400517.4:2000SEP08	2492	2572	forward 2	TM	N out
125	LI:400517.4:2000SEP08	2861	2944	forward 2	TM	N out
125	LI:400517.4:2000SEP08	864	944	forward 3	TM	
125	LI:400517.4:2000SEP08	1149	1229	forward 3	TM	
125	LI:400517.4:2000SEP08	2862	2936	forward 3	TM	
127	LI:1012560.1:2000SEP08	403	486	forward 1	TM	N out
127	LI:1012560.1:2000SEP08	784	870	forward 1	TM	N out
127	LI:1012560.1:2000SEP08	239	319	forward 2	TM	N out
127	LI:1012560.1:2000SEP08	695	781	forward 2	TM	N out
127	LI:1012560.1:2000SEP08	645	731	forward 3	TM	N in
127	LI:1012560.1:2000SEP08	804	887	forward 3	TM	N in
128	LI:427997.4:2000SEP08	148	222	forward 1	TM	N out
128	LI:427997.4:2000SEP08	748	816	forward 1	TM	N out
128	LI:427997.4:2000SEP08	1039	1125	forward 1	TM	N out
128	LI:427997.4:2000SEP08	1642	1728	forward 1	TM	N out
128	LI:427997.4:2000SEP08	1759	1833	forward 1	TM	N out
128	LI:427997.4:2000SEP08	134	220	forward 2	TM	N in
128	LI:427997.4:2000SEP08	737	823	forward 2	TM	N in
128	LI:427997.4:2000SEP08	1205	1291	forward 2	TM	N in
128	LI:427997.4:2000SEP08	1364	1447	forward 2	TM	N in
128	LI:427997.4:2000SEP08	1463	1531	forward 2	TM	N in
128	LI:427997.4:2000SEP08	1772	1858	forward 2	TM	N in
128	LI:427997.4:2000SEP08	150	236	forward 3	TM	N in
128	LI:427997.4:2000SEP08	672	755	forward 3	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
128	LI:427997.4:2000SEP08	786	869	forward 3	TM	N in
128	LI:427997.4:2000SEP08	1665	1751	forward 3	TM	N in
128	LI:427997.4:2000SEP08	1788	1874	forward 3	TM	N in
128	LI:427997.4:2000SEP08	2169	2249	forward 3	TM	N in
129	LI:197899.1:2000SEP08	205	291	forward 1	TM	N in
129	LI:197899.1:2000SEP08	301	381	forward 1	TM	N in
129	LI:197899.1:2000SEP08	439	513	forward 1	TM	N in
129	LI:197899.1:2000SEP08	604	690	forward 1	TM	N in
129	LI:197899.1:2000SEP08	688	774	forward 1	TM	N in
129	LI:197899.1:2000SEP08	988	1044	forward 1	TM	N in
129	LI:197899.1:2000SEP08	1525	1584	forward 1	TM	N in
129	LI:197899.1:2000SEP08	2029	2115	forward 1	TM	N in
129	LI:197899.1:2000SEP08	197	283	forward 2	TM	N out
129	LI:197899.1:2000SEP08	326	412	forward 2	TM	N out
129	LI:197899.1:2000SEP08	824	910	forward 2	TM	N out
129	LI:197899.1:2000SEP08	956	1042	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1094	1180	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1241	1303	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1328	1390	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1529	1615	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1628	1702	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1727	1789	forward 2	TM	N out
129	LI:197899.1:2000SEP08	1805	1867	forward 2	TM	N out
129	LI:197899.1:2000SEP08	2075	2140	forward 2	TM	N out
129	LI:197899.1:2000SEP08	90	140	forward 3	TM	
129	LI:197899.1:2000SEP08	321	407	forward 3	TM	
129	LI:197899.1:2000SEP08	951	1016	forward 3	TM	
129	LI:197899.1:2000SEP08	1059	1133	forward 3	TM	
129	LI:197899.1:2000SEP08	1485	1568	forward 3	TM	
130	LG:334199.1:2000SEP08	80	154	forward 2	TM	N out
131	LG:334345.1:2000SEP08	503	586	forward 2	TM	N in
131	LG:334345.1:2000SEP08	60	125	forward 3	TM	N in
132	LG:228092.1:2000SEP08	319	402	forward 1	TM	N in
132	LG:228092.1:2000SEP08	511	597	forward 1	TM	N in
132	LG:228092.1:2000SEP08	818	877	forward 2	TM	
132	LG:228092.1:2000SEP08	333	419	forward 3	TM	N in
132	LG:228092.1:2000SEP08	1113	1184	forward 3	TM	N in
133	LG:098580.1:2000SEP08	520	606	forward 1	TM	N out
133	LG:098580.1:2000SEP08	640	708	forward 1	TM	N out
133	LG:098580.1:2000SEP08	494	580	forward 2	TM	N out
133	LG:098580.1:2000SEP08	668	739	forward 2	TM	N out
133	LG:098580.1:2000SEP08	513	599	forward 3	TM	N out
133	LG:098580.1:2000SEP08	636	722	forward 3	TM	N out
134	LG:969572.1:2000SEP08	55	141	forward 1	TM	N out
134	LG:969572.1:2000SEP08	256	342	forward 1	TM	N out
135	LG:196958.1:2000SEP08	287	367	forward 2	TM	N in
135	LG:196958.1:2000SEP08	1007	1057	forward 2	TM	N in
135	LG:196958.1:2000SEP08	1247	1333	forward 2	TM	N in
135	LG:196958.1:2000SEP08	1281	1367	forward 3	TM	N out



TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
136	LG:1087811.1:2000SEP08	799	855	forward 1	TM	N out
136	LG:1087811.1:2000SEP08	728	814	forward 2	TM	N out
137	LG:1327885.1:2000SEP08	84	167	forward 3	TM	N in
137	LG:1327885.1:2000SEP08	471	548	forward 3	TM	N in
139	LI:897616.1:2000SEP08	63	149	forward 3	TM	N out
141	LI:027066.6:2000SEP08	707	793	forward 2	TM	N out
142	LI:1074263.1:2000SEP08	21	101	forward 3	TM	N out
143	LI:334345.1:2000SEP08	592	675	forward 1	TM	N in
143	LI:334345.1:2000SEP08	128	202	forward 2	TM	N out
144	LI:1093914.1:2000SEP08	772	825	forward 1	TM	N out
144	LI:1093914.1:2000SEP08	479	550	forward 2	TM	N out
144	LI:1093914.1:2000SEP08	324	374	forward 3	TM	N in
145	LI:1188168.1:2000SEP08	366	449	forward 3	TM	N in
145	LI:1188168.1:2000SEP08	2598	2645	forward 3	TM	N in
146	LI:1065168.1:2000SEP08	295	381	forward 1	TM	N in
146	LI:1065168.1:2000SEP08	194	280	forward 2	TM	N out
147	LI:1180418.1:2000SEP08	799	855	forward 1	TM	N out
149	LG:1078420.1:2000SEP08	707	784	forward 2	TM	N out
149	LG:1078420.1:2000SEP08	669	749	forward 3	TM	N in
150	LG:1397599.1:2000SEP08	46	120	forward 1	TM	N out
150	LG:1397599.1:2000SEP08	178	261	forward 1	TM	N out
150	LG:1397599.1:2000SEP08	29	115	forward 2	TM	N out
150	LG:1397599.1:2000SEP08	149	232	forward 2	TM	N out
150	LG:1397599.1:2000SEP08	24	110	forward 3	TM	N in
150	LG:1397599.1:2000SEP08	141	191	forward 3	TM	N in
151	LG:1397655.2:2000SEP08	406	489	forward 1	TM	N out
151	LG:1397655.2:2000SEP08	131	190	forward 2	TM	N out
151	LG:1397655.2:2000SEP08	401	451	forward 2	TM	N out
151	LG:1397655.2:2000SEP08	444	527	forward 3	TM	N out
152	LG:241055.1:2000SEP08	19	90	forward 1	TM	N out
152	LG:241055.1:2000SEP08	172	255	forward 1	TM	N out
152	LG:241055.1:2000SEP08	1045	1110	forward 1	TM	N out
152	LG:241055.1:2000SEP08	51	113	forward 3	TM	N out
152	LG:241055.1:2000SEP08	162	224	forward 3	TM	N out
153	LG:1101065.1:2000SEP08	1	57	forward 1	TM	N out
153	LG:1101065.1:2000SEP08	11	73	forward 2	TM	N in
153	LG:1101065.1:2000SEP08	92	139	forward 2	TM	N in
153	LG:1101065.1:2000SEP08	732	809	forward 3	TM	N out
155	LI:348991.1:2000SEP08	835	900	forward 1	TM	N in
155	LI:348991.1:2000SEP08	803	889	forward 2	TM	N out
155	LI:348991.1:2000SEP08	843	899	forward 3	TM	N out
158	LI:815686.1:2000SEP08	163	246	forward 1	TM	N in
158	LI:815686.1:2000SEP08	673	759	forward 1	TM	N in
158	LI:815686.1:2000SEP08	868	954	forward 1	TM	N in
158	LI:815686.1:2000SEP08	1108	1173	forward 1	TM	N in
158	LI:815686.1:2000SEP08	1201	1287	forward 1	TM	N in
158	LI:815686.1:2000SEP08	1369	1452	forward 1	TM	N in
158	LI:815686.1:2000SEP08	1597	1659	forward 1	TM	N in
158	LI:815686.1:2000SEP08	164	250	forward 2	TM	N out

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
158	LI:815686.1:2000SEP08	662	748	forward 2	TM	N out
158	LI:815686.1:2000SEP08	824	910	forward 2	TM	N out
158	LI:815686.1:2000SEP08	938	988	forward 2	TM	N out
158	LI:815686.1:2000SEP08	1031	1117	forward 2	TM	N out
158	LI:815686.1:2000SEP08	1118	1192	forward 2	TM	N out
158	LI:815686.1:2000SEP08	1559	1645	forward 2	TM	N out
158	LI:815686.1:2000SEP08	12	74	forward 3	TM	N in
158	LI:815686.1:2000SEP08	474	557	forward 3	TM	N in
158	LI:815686.1:2000SEP08	813	899	forward 3	TM	N in
158	LI:815686.1:2000SEP08	1110	1196	forward 3	TM	N in
158	LI:815686.1:2000SEP08	1389	1451	forward 3	TM	N in
158	LI:815686.1:2000SEP08	1470	1532	forward 3	TM	N in
159	LI:1167327.2:2000SEP08	31	117	forward 1	TM	N out
159	LI:1167327.2:2000SEP08	151	228	forward 1	TM	N out
159	LI:1167327.2:2000SEP08	23	109	forward 2	TM	N out
159	LI:1167327.2:2000SEP08	263	349	forward 2	TM	N out
159	LI:1167327.2:2000SEP08	48	122	forward 3	TM	N out
159	LI:1167327.2:2000SEP08	150	236	forward 3	TM	N out
161	LG:331593.1:2000SEP08	902	973	forward 2	TM	N out
161	LG:331593.1:2000SEP08	579	656	forward 3	TM	N out
162	LI:1094174.1:2000SEP08	1270	1329	forward 1	TM	N out
162	LI:1094174.1:2000SEP08	1656	1742	forward 3	TM	N in
163	LI:814362.1:2000SEP08	331	417	forward 1	TM	N out
163	LI:814362.1:2000SEP08	383	469	forward 2	TM	N out
163	LI:814362.1:2000SEP08	18	95	forward 3	TM	N in
164	LI:219542.1:2000SEP08	13	72	forward 1	TM	N out
164	LI:219542.1:2000SEP08	265	345	forward 1	TM	N out
164	LI:219542.1:2000SEP08	279	365	forward 3	TM	N in
165	LI:726197.1:2000SEP08	187	267	forward 1	TM	N in
166	LI:1075314.1:2000SEP08	235	297	forward 1	TM	N in
166	LI:1075314.1:2000SEP08	325	387	forward 1	TM	N in
166	LI:1075314.1:2000SEP08	580	642	forward 1	TM	N in
166	LI:1075314.1:2000SEP08	694	765	forward 1	TM	N in
166	LI:1075314.1:2000SEP08	422	484	forward 2	TM	N out
166	LI:1075314.1:2000SEP08	509	571	forward 2	TM	N out
166	LI:1075314.1:2000SEP08	657	737	forward 3	TM	N out
168	LG:336265.1:2000SEP08	868	930	forward 1	TM	N out
168	LG:336265.1:2000SEP08	943	1005	forward 1	TM	N out
168	LG:336265.1:2000SEP08	848	919	forward 2	TM	N out
168	LG:336265.1:2000SEP08	965	1024	forward 2	TM	N out
168	LG:336265.1:2000SEP08	825	911	forward 3	TM	N in
168	LG:336265.1:2000SEP08	951	1013	forward 3	TM	N in
168	LG:336265.1:2000SEP08	1038	1100	forward 3	TM	N in
168	LG:336265.1:2000SEP08	1497	1583	forward 3	TM	N in
169	LG:407788.2:2000SEP08	253	336	forward 1	TM	N out
169	LG:407788.2:2000SEP08	493	558	forward 1	TM	N out
169	LG:407788.2:2000SEP08	562	648	forward 1	TM	N out
169	LG:407788.2:2000SEP08	278	355	forward 2	TM	
169	LG:407788.2:2000SEP08	542	604	forward 2	TM	

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
169	LG:407788.2:2000SEP08	626	688	forward 2	TM	
169	LG:407788.2:2000SEP08	24	110	forward 3	TM	N out
169	LG:407788.2:2000SEP08	282	365	forward 3	TM	N out
169	LG:407788.2:2000SEP08	612	686	forward 3	TM	N out
170	LG:1326925.1:2000SEP08	62	142	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	164	226	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	242	310	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	320	388	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	479	541	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	560	622	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	740	826	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	890	961	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	959	1009	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	1103	1165	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	1178	1240	forward 2	TM	N out
170	LG:1326925.1:2000SEP08	1277	1333	forward 2	TM	N out
171	LI:332655.2:2000SEP08	541	621	forward 1	TM	N out
171	LI:332655.2:2000SEP08	536	598	forward 2	TM	N out
171	LI:332655.2:2000SEP08	626	688	forward 2	TM	N out
171	LI:332655.2:2000SEP08	887	949	forward 2	TM	N out
171	LI:332655.2:2000SEP08	965	1027	forward 2	TM	N out
171	LI:332655.2:2000SEP08	912	998	forward 3	TM	N in
172	LI:1184621.4:2000SEP08	265	351	forward 1	TM	N out
172	LI:1184621.4:2000SEP08	499	564	forward 1	TM	N out
172	LI:1184621.4:2000SEP08	568	654	forward 1	TM	N out
172	LI:1184621.4:2000SEP08	847	933	forward 1	TM	N out
172	LI:1184621.4:2000SEP08	284	370	forward 2	TM	N out
172	LI:1184621.4:2000SEP08	554	616	forward 2	TM	N out
172	LI:1184621.4:2000SEP08	638	700	forward 2	TM	N out
172	LI:1184621.4:2000SEP08	39	89	forward 3	TM	N out
172	LI:1184621.4:2000SEP08	288	365	forward 3	TM	N out
172	LI:1184621.4:2000SEP08	612	686	forward 3	TM	N out
172	LI:1184621.4:2000SEP08	804	890	forward 3	TM	N out
173	LI:2051386.1:2000SEP08	272	358	forward 2	TM	N out
173	LI:2051386.1:2000SEP08	809	886	forward 2	TM	N out
173	LI:2051386.1:2000SEP08	441	503	forward 3	TM	N in
173	LI:2051386.1:2000SEP08	525	587	forward 3	TM	N in
173	LI:2051386.1:2000SEP08	609	671	forward 3	TM	N in
173	LI:2051386.1:2000SEP08	693	755	forward 3	TM	N in
174	LG:362757.1:2000SEP08	506	571	forward 2	TM	N out
174	LG:362757.1:2000SEP08	450	524	forward 3	TM	N out
175	LG:406770.1:2000SEP08	1436	1510	forward 2	TM	N in
177	LG:001929.1:2000SEP08	125	205	forward 2	TM	N out
177	LG:001929.1:2000SEP08	1469	1543	forward 2	TM	N out
177	LG:001929.1:2000SEP08	1416	1493	forward 3	TM	N in
177	LG:001929.1:2000SEP08	1665	1733	forward 3	TM	N in
178	LI:401322.1:2000SEP08	332	385	forward 2	TM	N out
179	LI:208748.1:2000SEP08	673	726	forward 1	TM	N out
179	LI:208748.1:2000SEP08	724	807	forward 1	TM	N out

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
179	LI:208748.1:2000SEP08	2152	2229	forward 1	TM	N out
179	LI:208748.1:2000SEP08	98	184	forward 2	TM	N out
179	LI:208748.1:2000SEP08	320	373	forward 2	TM	N out
179	LI:208748.1:2000SEP08	794	880	forward 2	TM	N out
179	LI:208748.1:2000SEP08	941	1006	forward 2	TM	N out
179	LI:208748.1:2000SEP08	1346	1393	forward 2	TM	N out
179	LI:208748.1:2000SEP08	66	152	forward 3	TM	N in
179	LI:208748.1:2000SEP08	516	575	forward 3	TM	N in
179	LI:208748.1:2000SEP08	783	839	forward 3	TM	N in
179	LI:208748.1:2000SEP08	1080	1166	forward 3	TM	N in
179	LI:208748.1:2000SEP08	1335	1391	forward 3	TM	N in
179	LI:208748.1:2000SEP08	1647	1724	forward 3	TM	N in
179	LI:208748.1:2000SEP08	1896	1964	forward 3	TM	N in
179	LI:208748.1:2000SEP08	2271	2357	forward 3	TM	N in
180	LI:407242.1:2000SEP08	226	288	forward 1	TM	N in
180	LI:407242.1:2000SEP08	514	573	forward 1	TM	N in
180	LI:407242.1:2000SEP08	1774	1836	forward 1	TM	N in
180	LI:407242.1:2000SEP08	1355	1441	forward 2	TM	N in
180	LI:407242.1:2000SEP08	255	341	forward 3	TM	N in
180	LI:407242.1:2000SEP08	387	440	forward 3	TM	N in
180	LI:407242.1:2000SEP08	510	581	forward 3	TM	N in
180	LI:407242.1:2000SEP08	747	821	forward 3	TM	N in
180	LI:407242.1:2000SEP08	1659	1745	forward 3	TM	N in
181	LI:403409.1:2000SEP08	136	222	forward 1	TM	N out
181	LI:403409.1:2000SEP08	973	1029	forward 1	TM	N out
181	LI:403409.1:2000SEP08	1285	1371	forward 1	TM	N out
181	LI:403409.1:2000SEP08	182	268	forward 2	TM	N in
182	LI:450798.1:2000SEP08	655	723	forward 1	TM	N in
182	LI:450798.1:2000SEP08	1531	1617	forward 1	TM	N in
182	LI:450798.1:2000SEP08	911	979	forward 2	TM	N out
182	LI:450798.1:2000SEP08	1535	1603	forward 2	TM	N out
183	LI:410317.1:2000SEP08	1201	1272	forward 1	TM	N out
183	LI:410317.1:2000SEP08	1097	1183	forward 2	TM	N in
184	LI:340268.1:2000SEP08	136	192	forward 1	TM	N out
185	LI:2051671.1:2000SEP08	307	372	forward 1	TM	N out
185	LI:2051671.1:2000SEP08	554	640	forward 2	TM	N in
186	LG:998844.1:2000SEP08	189	263	forward 3	TM	N out
187	LG:1043787.1:2000SEP08	22	108	forward 1	TM	N in
187	LG:1043787.1:2000SEP08	23	85	forward 2	TM	N out
188	LG:1098931.16:2000SEP08	241	327	forward 1	TM	N in
188	LG:1098931.16:2000SEP08	218	298	forward 2	TM	N in
189	LG:199423.2:2000SEP08	19	90	forward 1	TM	N out
189	LG:199423.2:2000SEP08	115	183	forward 1	TM	N out
190	LI:1075297.1:2000SEP08	83	157	forward 2	TM	N out
190	LI:1075297.1:2000SEP08	287	343	forward 2	TM	N out
190	LI:1075297.1:2000SEP08	344	409	forward 2	TM	N out
190	LI:1075297.1:2000SEP08	624	710	forward 3	TM	N in
191	LI:1043321.1:2000SEP08	436	498	forward 1	TM	N in
191	LI:1043321.1:2000SEP08	523	585	forward 1	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
191	U:1043321.1:2000SEP08	159	239	forward 3	TM	N in
192	U:297070.1:2000SEP08	583	648	forward 1	TM	
192	U:297070.1:2000SEP08	575	661	forward 2	TM	N out
192	U:297070.1:2000SEP08	12	62	forward 3	TM	N out
193	U:1085041.1:2000SEP08	54	116	forward 3	TM	
194	U:1071544.1:2000SEP08	425	511	forward 2	TM	N in
195	U:2052480.1:2000SEP08	245	307	forward 2	TM	N out
195	U:2052480.1:2000SEP08	1079	1165	forward 2	TM	N out
195	U:2052480.1:2000SEP08	12	77	forward 3	TM	N out
197	LG:450581.1:2000SEP08	274	351	forward 1	TM	
197	LG:450581.1:2000SEP08	376	426	forward 1	TM	
197	LG:450581.1:2000SEP08	239	301	forward 2	TM	N out
197	LG:450581.1:2000SEP08	317	379	forward 2	TM	N out
200	LG:452089.1:2000SEP08	694	771	forward 1	TM	N out
200	LG:452089.1:2000SEP08	713	769	forward 2	TM	N out
201	LG:1099416.1:2000SEP08	622	708	forward 1	TM	N out
201	LG:1099416.1:2000SEP08	596	682	forward 2	TM	N out
201	LG:1099416.1:2000SEP08	588	674	forward 3	TM	N out
202	LG:255713.1:2000SEP08	351	413	forward 3	TM	N in
203	LG:998903.1:2000SEP08	493	579	forward 1	TM	N out
203	LG:998903.1:2000SEP08	269	355	forward 2	TM	N out
203	LG:998903.1:2000SEP08	402	488	forward 3	TM	
204	LG:1119656.1:2000SEP08	52	120	forward 1	TM	N out
204	LG:1119656.1:2000SEP08	169	243	forward 1	TM	N out
204	LG:1119656.1:2000SEP08	26	100	forward 2	TM	N in
204	LG:1119656.1:2000SEP08	194	271	forward 2	TM	N in
204	LG:1119656.1:2000SEP08	36	107	forward 3	TM	N in
204	LG:1119656.1:2000SEP08	153	239	forward 3	TM	N in
206	LG:1323741.1:2000SEP08	23	109	forward 2	TM	N out
206	LG:1323741.1:2000SEP08	12	71	forward 3	TM	N out
209	LG:1097562.1:2000SEP08	48	122	forward 3	TM	N out
210	LG:998868.1:2000SEP08	694	753	forward 1	TM	N in
210	LG:998868.1:2000SEP08	728	814	forward 2	TM	N out
211	LG:1063383.1:2000SEP08	238	309	forward 1	TM	N in
211	LG:1063383.1:2000SEP08	505	579	forward 1	TM	N in
211	LG:1063383.1:2000SEP08	236	322	forward 2	TM	N in
211	LG:1063383.1:2000SEP08	462	548	forward 3	TM	N in
213	U:449404.1:2000SEP08	163	222	forward 1	TM	N out
213	U:449404.1:2000SEP08	523	594	forward 1	TM	N out
213	U:449404.1:2000SEP08	200	280	forward 2	TM	N out
214	U:449941.2:2000SEP08	173	259	forward 2	TM	N out
215	U:450229.1:2000SEP08	374	454	forward 2	TM	N in
216	U:450399.3:2000SEP08	145	231	forward 1	TM	N out
216	U:450399.3:2000SEP08	544	597	forward 1	TM	N out
216	U:450399.3:2000SEP08	491	565	forward 2	TM	N out
216	U:450399.3:2000SEP08	417	503	forward 3	TM	N in
217	U:455771.1:2000SEP08	199	285	forward 1	TM	N out
220	U:728055.1:2000SEP08	205	291	forward 1	TM	N out
223	U:1084329.1:2000SEP08	357	410	forward 3	TM	N in

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
225	LI:1086066.1:2000SEP08	99	185	forward 3	TM	N out
226	LI:223142.1:2000SEP08	826	888	forward 1	TM	
226	LI:223142.1:2000SEP08	973	1059	forward 1	TM	
226	LI:223142.1:2000SEP08	770	841	forward 2	TM	N in
226	LI:223142.1:2000SEP08	995	1051	forward 2	TM	N in
226	LI:223142.1:2000SEP08	798	857	forward 3	TM	N out
227	LI:885368.1:2000SEP08	540	608	forward 3	TM	N in
230	LI:449413.2:2000SEP08	349	435	forward 1	TM	N out
231	LI:450105.1:2000SEP08	208	294	forward 1	TM	N in
233	LI:1142855.1:2000SEP08	17	91	forward 2	TM	N out
233	LI:1142855.1:2000SEP08	245	295	forward 2	TM	N out
233	LI:1142855.1:2000SEP08	344	430	forward 2	TM	N out
233	LI:1142855.1:2000SEP08	51	98	forward 3	TM	N out
235	LI:817845.1:2000SEP08	460	513	forward 1	TM	N in
235	LI:817845.1:2000SEP08	515	601	forward 2	TM	N in
237	LI:815874.1:2000SEP08	467	550	forward 2	TM	N in
237	LI:815874.1:2000SEP08	513	587	forward 3	TM	N out
238	LI:255713.1:2000SEP08	621	674	forward 3	TM	
239	LI:035973.1:2000SEP08	790	849	forward 1	TM	N in
239	LI:035973.1:2000SEP08	626	712	forward 2	TM	N out
239	LI:035973.1:2000SEP08	642	728	forward 3	TM	N in
240	LI:1138110.1:2000SEP08	22	108	forward 1	TM	N in
240	LI:1138110.1:2000SEP08	154	240	forward 1	TM	N in
240	LI:1138110.1:2000SEP08	47	121	forward 2	TM	N out
240	LI:1138110.1:2000SEP08	170	244	forward 2	TM	N out
240	LI:1138110.1:2000SEP08	51	110	forward 3	TM	N in
240	LI:1138110.1:2000SEP08	195	272	forward 3	TM	N in
242	LI:1092460.1:2000SEP08	181	243	forward 1	TM	N out
243	LI:399421.1:2000SEP08	310	396	forward 1	TM	N in
243	LI:399421.1:2000SEP08	1681	1767	forward 1	TM	N in
243	LI:399421.1:2000SEP08	1900	1950	forward 1	TM	N in
243	LI:399421.1:2000SEP08	59	112	forward 2	TM	N out
243	LI:399421.1:2000SEP08	593	664	forward 2	TM	N out
243	LI:399421.1:2000SEP08	797	853	forward 2	TM	N out
243	LI:399421.1:2000SEP08	1445	1519	forward 2	TM	N out
243	LI:399421.1:2000SEP08	1640	1705	forward 2	TM	N out
243	LI:399421.1:2000SEP08	2000	2074	forward 2	TM	N out
243	LI:399421.1:2000SEP08	666	752	forward 3	TM	N out
243	LI:399421.1:2000SEP08	1461	1517	forward 3	TM	N out
243	LI:399421.1:2000SEP08	1893	1955	forward 3	TM	N out
244	LI:816655.2:2000SEP08	373	435	forward 1	TM	N in
244	LI:816655.2:2000SEP08	466	528	forward 1	TM	N in
244	LI:816655.2:2000SEP08	871	948	forward 1	TM	N in
244	LI:816655.2:2000SEP08	1099	1185	forward 1	TM	N in
244	LI:816655.2:2000SEP08	344	427	forward 2	TM	N in
244	LI:816655.2:2000SEP08	1127	1213	forward 2	TM	N in
244	LI:816655.2:2000SEP08	453	539	forward 3	TM	N in
244	LI:816655.2:2000SEP08	1062	1127	forward 3	TM	N in
245	LG:414732.1:2000SEP08	40	93	forward 1	TM	N out

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
245	LG:414732.1:2000SEP08	156	233	forward 3	TM	N out
246	LG:1140250.1:2000SEP08	410	493	forward 2	TM	N in
247	LG:174022.1:2000SEP08	817	894	forward 1	TM	N in
247	LG:174022.1:2000SEP08	671	757	forward 2	TM	N in
247	LG:174022.1:2000SEP08	797	883	forward 2	TM	N in
248	LI:002811.1:2000SEP08	313	384	forward 1	TM	N out
248	LI:002811.1:2000SEP08	520	573	forward 1	TM	N out
248	LI:002811.1:2000SEP08	595	654	forward 1	TM	N out
248	LI:002811.1:2000SEP08	248	322	forward 2	TM	N out
248	LI:002811.1:2000SEP08	282	368	forward 3	TM	N in
248	LI:002811.1:2000SEP08	465	551	forward 3	TM	N in
248	LI:002811.1:2000SEP08	567	653	forward 3	TM	N in
249	LI:414732.2:2000SEP08	34	93	forward 1	TM	N out
249	LI:414732.2:2000SEP08	24	110	forward 3	TM	N out
249	LI:414732.2:2000SEP08	159	236	forward 3	TM	N out
250	LI:1019920.1:2000SEP08	360	422	forward 3	TM	N in
251	LI:1038336.1:2000SEP08	53	115	forward 2	TM	N out
251	LI:1038336.1:2000SEP08	128	190	forward 2	TM	N out
252	LI:1177772.11:2000SEP08	49	135	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	169	237	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	1105	1191	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	1222	1308	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	1333	1419	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	1561	1647	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	1705	1758	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	2152	2238	forward 1	TM	N out
252	LI:1177772.11:2000SEP08	1082	1144	forward 2	TM	N in
252	LI:1177772.11:2000SEP08	1157	1219	forward 2	TM	N in
252	LI:1177772.11:2000SEP08	1226	1306	forward 2	TM	N in
252	LI:1177772.11:2000SEP08	1535	1609	forward 2	TM	N in
252	LI:1177772.11:2000SEP08	36	98	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	135	197	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	234	296	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	1053	1139	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	1188	1265	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	1491	1577	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	1866	1952	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	2004	2066	forward 3	TM	N out
252	LI:1177772.11:2000SEP08	2091	2153	forward 3	TM	N out
253	LI:205642.2:2000SEP08	178	261	forward 1	TM	N out
253	LI:205642.2:2000SEP08	541	591	forward 1	TM	N out
253	LI:205642.2:2000SEP08	521	577	forward 2	TM	N out
254	LG:449685.1:2000SEP08	491	565	forward 2	TM	N in
254	LG:449685.1:2000SEP08	132	206	forward 3	TM	N in
254	LG:449685.1:2000SEP08	447	512	forward 3	TM	N in
255	LG:453922.1:2000SEP08	465	539	forward 3	TM	N out
256	LG:476342.3:2000SEP08	70	153	forward 1	TM	N out
257	LI:336801.1:2000SEP08	613	699	forward 1	TM	N out
257	LI:336801.1:2000SEP08	611	697	forward 2	TM	N out

TABLE 4

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain	Topology
258	LI:449685.1:2000SEP08	499	552	forward 1	TM	
258	LI:449685.1:2000SEP08	446	511	forward 2	TM	N out
258	LI:449685.1:2000SEP08	132	218	forward 3	TM	N in
259	LI:476342.1:2000SEP08	39	122	forward 3	TM	N out
260	LI:1072804.1:2000SEP08	500	586	forward 2	TM	N out
260	LI:1072804.1:2000SEP08	276	332	forward 3	TM	N out
261	LI:455450.1:2000SEP08	422	490	forward 2	TM	N out
263	LI:1013729.1:2000SEP08	475	540	forward 1	TM	N in
263	LI:1013729.1:2000SEP08	470	556	forward 2	TM	N out
263	LI:1013729.1:2000SEP08	507	554	forward 3	TM	N in
264	LI:2050322.2:2000SEP08	1103	1156	forward 2	TM	N out
266	LI:2053076.1:2000SEP08	112	198	forward 1	TM	N in
266	LI:2053076.1:2000SEP08	490	576	forward 1	TM	N in
266	LI:2053076.1:2000SEP08	62	148	forward 2	TM	N in
266	LI:2053076.1:2000SEP08	720	806	forward 3	TM	
268	LG:406709.1:2000SEP08	183	233	forward 3	TM	N in
269	LG:347863.9:2000SEP08	125	211	forward 2	TM	N out
269	LG:347863.9:2000SEP08	497	562	forward 2	TM	N out
269	LG:347863.9:2000SEP08	24	110	forward 3	TM	N out
269	LG:347863.9:2000SEP08	156	230	forward 3	TM	N out
271	LI:347635.1:2000SEP08	664	735	forward 1	TM	N in
271	LI:347635.1:2000SEP08	1468	1554	forward 1	TM	N in
271	LI:347635.1:2000SEP08	815	901	forward 2	TM	N out
271	LI:347635.1:2000SEP08	1547	1597	forward 2	TM	N out
271	LI:347635.1:2000SEP08	414	488	forward 3	TM	N out
271	LI:347635.1:2000SEP08	600	650	forward 3	TM	N out
272	LI:013685.1:2000SEP08	22	108	forward 1	TM	N out
272	LI:013685.1:2000SEP08	1483	1536	forward 1	TM	N out
272	LI:013685.1:2000SEP08	221	304	forward 2	TM	N out
272	LI:013685.1:2000SEP08	653	700	forward 2	TM	N out
272	LI:013685.1:2000SEP08	711	767	forward 3	TM	N in
273	LI:406709.1:2000SEP08	183	233	forward 3	TM	N in
274	LI:2052938.1:2000SEP08	437	505	forward 2	TM	



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
1	LG:405741.3:2000SEP08	g4076447	772	1210
1	LG:405741.3:2000SEP08	g6473689	757	1199
1	LG:405741.3:2000SEP08	g764385	877	1199
1	LG:405741.3:2000SEP08	g4332695	906	1198
1	LG:405741.3:2000SEP08	g5742159	755	1198
1	LG:405741.3:2000SEP08	g4326577	1008	1198
1	LG:405741.3:2000SEP08	g1721162	906	1197
1	LG:405741.3:2000SEP08	g4393490	746	1196
1	LG:405741.3:2000SEP08	g4004428	793	1196
1	LG:405741.3:2000SEP08	g6658709	732	1196
1	LG:405741.3:2000SEP08	g4510098	936	1196
1	LG:405741.3:2000SEP08	g2224160	838	1196
1	LG:405741.3:2000SEP08	g1391615	774	1193
1	LG:405741.3:2000SEP08	2605255H1	953	1194
1	LG:405741.3:2000SEP08	g2986591	948	1193
1	LG:405741.3:2000SEP08	1632154H1	985	1193
1	LG:405741.3:2000SEP08	g4327028	998	1193
1	LG:405741.3:2000SEP08	g1225424	827	1193
1	LG:405741.3:2000SEP08	g2705577	772	1193
1	LG:405741.3:2000SEP08	g6710344	797	1193
1	LG:405741.3:2000SEP08	2888488H1	1072	1193
1	LG:405741.3:2000SEP08	2605255F6	953	1193
1	LG:405741.3:2000SEP08	238539R6	733	1185
1	LG:405741.3:2000SEP08	g1080925	861	1180
1	LG:405741.3:2000SEP08	7624412H1	973	1169
1	LG:405741.3:2000SEP08	7402118H1	725	1164
1	LG:405741.3:2000SEP08	614864T6	629	1158
1	LG:405741.3:2000SEP08	2605255T6	953	1156
1	LG:405741.3:2000SEP08	1437574T6	823	1155
1	LG:405741.3:2000SEP08	3696915T6	1057	1149
1	LG:405741.3:2000SEP08	1501410T6	666	1143
1	LG:405741.3:2000SEP08	2779559T6	711	1134
1	LG:405741.3:2000SEP08	6848176H1	599	1115
1	LG:405741.3:2000SEP08	7624412J1	934	1097
1	LG:405741.3:2000SEP08	g4310285	632	1075
1	LG:405741.3:2000SEP08	g3917108	632	1019
1	LG:405741.3:2000SEP08	614864R6	481	960
1	LG:405741.3:2000SEP08	345217H1	759	955
1	LG:405741.3:2000SEP08	1581267F6	479	953
1	LG:405741.3:2000SEP08	238539H1	733	945
1	LG:405741.3:2000SEP08	6822694J1	406	888
1	LG:405741.3:2000SEP08	6822694H1	406	888
1	LG:405741.3:2000SEP08	g1272796	337	789
1	LG:405741.3:2000SEP08	614864H1	481	725
1	LG:405741.3:2000SEP08	5655089H1	246	715
1	LG:405741.3:2000SEP08	1581267H1	479	693
1	LG:405741.3:2000SEP08	5327332H1	411	696
1	LG:405741.3:2000SEP08	1581267T6	600	684
1	LG:405741.3:2000SEP08	1504659T1	571	684

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
1	LG:405741.3:2000SEP08	g1401461	347	680
1	LG:405741.3:2000SEP08	5327610H1	411	638
1	LG:405741.3:2000SEP08	g1389301	265	633
1	LG:405741.3:2000SEP08	g1081018	265	634
1	LG:405741.3:2000SEP08	g762007	337	609
1	LG:405741.3:2000SEP08	5392809H1	212	423
1	LG:405741.3:2000SEP08	4181142H1	153	417
1	LG:405741.3:2000SEP08	5288480F6	1	385
1	LG:405741.3:2000SEP08	5288480H1	1	256
2	LG:337194.1:2000SEP08	g2877617	1127	1490
2	LG:337194.1:2000SEP08	1391371H1	1238	1516
2	LG:337194.1:2000SEP08	3216293F6	1284	1847
2	LG:337194.1:2000SEP08	3216293H1	1284	1525
2	LG:337194.1:2000SEP08	5973325H1	1298	1517
2	LG:337194.1:2000SEP08	842273R1	1467	2031
2	LG:337194.1:2000SEP08	842273H1	1467	1724
2	LG:337194.1:2000SEP08	4758225H1	1475	1740
2	LG:337194.1:2000SEP08	1312387F6	1522	1859
2	LG:337194.1:2000SEP08	1312387H1	1522	1737
2	LG:337194.1:2000SEP08	999991T6	1636	2287
2	LG:337194.1:2000SEP08	1000054R1	1639	2093
2	LG:337194.1:2000SEP08	999991R6	1639	2146
2	LG:337194.1:2000SEP08	1000054H1	1639	1863
2	LG:337194.1:2000SEP08	6537357H1	1647	2055
2	LG:337194.1:2000SEP08	4351723H1	1689	2024
2	LG:337194.1:2000SEP08	6412057H1	1752	2160
2	LG:337194.1:2000SEP08	g1123225	1763	1843
2	LG:337194.1:2000SEP08	1312387T6	1795	2276
2	LG:337194.1:2000SEP08	6415837H1	1887	2175
2	LG:337194.1:2000SEP08	g4734612	1	439
2	LG:337194.1:2000SEP08	2673870F6	1	458
2	LG:337194.1:2000SEP08	2673870H1	1	224
2	LG:337194.1:2000SEP08	3040985H1	6	291
2	LG:337194.1:2000SEP08	6551863H1	37	235
2	LG:337194.1:2000SEP08	g5839202	51	395
2	LG:337194.1:2000SEP08	4213969F6	154	733
2	LG:337194.1:2000SEP08	4213969H1	154	402
2	LG:337194.1:2000SEP08	6339716H1	175	639
2	LG:337194.1:2000SEP08	6339716F6	175	772
2	LG:337194.1:2000SEP08	g2036487	191	444
2	LG:337194.1:2000SEP08	g2037866	216	512
2	LG:337194.1:2000SEP08	7762752H1	313	861
2	LG:337194.1:2000SEP08	g5450422	402	855
2	LG:337194.1:2000SEP08	g5639085	426	853
2	LG:337194.1:2000SEP08	g3678295	429	852
2	LG:337194.1:2000SEP08	g3924212	442	855
2	LG:337194.1:2000SEP08	g3678424	447	851
2	LG:337194.1:2000SEP08	3417306H1	485	725
2	LG:337194.1:2000SEP08	3417306F6	485	882

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
2	LG:337194.1:2000SEP08	2007764H1	507	716
2	LG:337194.1:2000SEP08	g3804340	549	854
2	LG:337194.1:2000SEP08	g4736790	551	850
2	LG:337194.1:2000SEP08	g2069693	565	1000
2	LG:337194.1:2000SEP08	5134964H1	567	842
2	LG:337194.1:2000SEP08	g3038708	691	852
2	LG:337194.1:2000SEP08	2500676H1	923	1170
2	LG:337194.1:2000SEP08	5134964T6	1896	2282
2	LG:337194.1:2000SEP08	6740988H1	1915	2324
2	LG:337194.1:2000SEP08	3216293T6	1950	2290
2	LG:337194.1:2000SEP08	3432816H1	1977	2218
2	LG:337194.1:2000SEP08	g4110765	1979	2324
2	LG:337194.1:2000SEP08	3246416T6	1999	2289
2	LG:337194.1:2000SEP08	3246416F6	2006	2324
2	LG:337194.1:2000SEP08	3246416H1	2007	2266
2	LG:337194.1:2000SEP08	2289722H1	2036	2264
2	LG:337194.1:2000SEP08	g2325587	2072	2325
2	LG:337194.1:2000SEP08	g2069302	2196	2332
3	LG:017108.4:2000SEP08	3642358F6	1	593
3	LG:017108.4:2000SEP08	3642358H1	1	295
3	LG:017108.4:2000SEP08	g2220930	19	247
3	LG:017108.4:2000SEP08	3531423H1	121	443
3	LG:017108.4:2000SEP08	373866H1	370	615
3	LG:017108.4:2000SEP08	311923H1	370	642
4	LG:372569.5:2000SEP08	3390549H1	18	323
4	LG:372569.5:2000SEP08	3390549F8	18	583
4	LG:372569.5:2000SEP08	2516619H1	11	320
4	LG:372569.5:2000SEP08	5886606H1	482	667
4	LG:372569.5:2000SEP08	2514054H1	500	741
4	LG:372569.5:2000SEP08	2512446H1	99	341
4	LG:372569.5:2000SEP08	5807337H1	481	673
4	LG:372569.5:2000SEP08	1644963F6	24	314
4	LG:372569.5:2000SEP08	1286148H1	23	188
4	LG:372569.5:2000SEP08	1644963H1	23	229
4	LG:372569.5:2000SEP08	7355708H1	24	463
4	LG:372569.5:2000SEP08	2515048H1	23	329
4	LG:372569.5:2000SEP08	3576172H1	750	1023
4	LG:372569.5:2000SEP08	4334360H1	665	939
4	LG:372569.5:2000SEP08	7392833H1	564	1092
4	LG:372569.5:2000SEP08	7601833H1	987	1396
4	LG:372569.5:2000SEP08	g4326942	1030	1389
4	LG:372569.5:2000SEP08	6758711H1	752	1327
4	LG:372569.5:2000SEP08	6758711J1	819	1398
4	LG:372569.5:2000SEP08	3576172F6	750	1288
4	LG:372569.5:2000SEP08	g1988216	1350	1531
4	LG:372569.5:2000SEP08	2014656T6	1388	1837
4	LG:372569.5:2000SEP08	g6039724	1400	1801
4	LG:372569.5:2000SEP08	2014656R6	1218	1709
4	LG:372569.5:2000SEP08	2013508H1	1218	1487

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
4	LG:372569.5:2000SEP08	2014656H1	1218	1484
4	LG:372569.5:2000SEP08	6820141J1	1254	1820
4	LG:372569.5:2000SEP08	g1313809	1077	1249
4	LG:372569.5:2000SEP08	g3873022	1076	1361
4	LG:372569.5:2000SEP08	g3429599	1146	1389
4	LG:372569.5:2000SEP08	7601833J1	1038	1395
4	LG:372569.5:2000SEP08	3752338H1	1	239
5	LG:968765.1:2000SEP08	5334376F8	197	709
5	LG:968765.1:2000SEP08	g1618573	1	112
5	LG:968765.1:2000SEP08	6796390F8	1	533
5	LG:968765.1:2000SEP08	6796390H1	1	240
5	LG:968765.1:2000SEP08	5334376H1	197	356
6	LG:255999.16:2000SEP08	7075546H1	1	533
6	LG:255999.16:2000SEP08	7075583H1	1	441
7	LG:977820.9:2000SEP08	7655023H1	1	578
7	LG:977820.9:2000SEP08	7655023J1	1	566
7	LG:977820.9:2000SEP08	g2835908	58	346
7	LG:977820.9:2000SEP08	3503804H1	139	451
7	LG:977820.9:2000SEP08	7644176J1	163	724
7	LG:977820.9:2000SEP08	6950721R8	213	815
7	LG:977820.9:2000SEP08	6867192H1	231	818
7	LG:977820.9:2000SEP08	6464824H1	609	1241
7	LG:977820.9:2000SEP08	6825706H1	904	1203
7	LG:977820.9:2000SEP08	6772076H1	911	1428
7	LG:977820.9:2000SEP08	1513539H1	1091	1293
7	LG:977820.9:2000SEP08	6935925H1	1108	1407
7	LG:977820.9:2000SEP08	723201H1	1114	1341
8	LI:1071608.1:2000SEP08	g1260446	2	316
8	LI:1071608.1:2000SEP08	6791379H1	1	397
8	LI:1071608.1:2000SEP08	g1614819	215	655
8	LI:1071608.1:2000SEP08	g1647514	244	543
9	LI:1074023.1:2000SEP08	6796546H1	1	475
9	LI:1074023.1:2000SEP08	6796546F8	1	510
9	LI:1074023.1:2000SEP08	6790876H1	9	540
9	LI:1074023.1:2000SEP08	6791780F8	9	487
9	LI:1074023.1:2000SEP08	6791780H1	9	539
9	LI:1074023.1:2000SEP08	6790876F8	9	619
9	LI:1074023.1:2000SEP08	6790876T8	16	673
9	LI:1074023.1:2000SEP08	6796546T8	175	651
10	LI:453570.1:2000SEP08	5911492T7	136	691
10	LI:453570.1:2000SEP08	5911492T9	305	681
10	LI:453570.1:2000SEP08	5911492T8	303	634
10	LI:453570.1:2000SEP08	5911492F8	1	467
10	LI:453570.1:2000SEP08	5911492F7	1	449
10	LI:453570.1:2000SEP08	5911492H1	1	271
11	LI:072072.1:2000SEP08	71678830V1	2203	2749
11	LI:072072.1:2000SEP08	71677535V1	2309	2797
11	LI:072072.1:2000SEP08	5306756H1	1174	1333
11	LI:072072.1:2000SEP08	5306856H1	1174	1341

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
11	LI:072072.1:2000SEP08	8112368H1	1180	1720
11	LI:072072.1:2000SEP08	g6837302	1185	1568
11	LI:072072.1:2000SEP08	g5634396	1273	1588
11	LI:072072.1:2000SEP08	g2167859	1292	1598
11	LI:072072.1:2000SEP08	71680895V1	2438	2886
11	LI:072072.1:2000SEP08	2531179H1	418	669
11	LI:072072.1:2000SEP08	g994114	2599	2896
11	LI:072072.1:2000SEP08	71677855V1	2653	3158
11	LI:072072.1:2000SEP08	71676834V1	2672	3256
11	LI:072072.1:2000SEP08	g4739904	2668	2904
11	LI:072072.1:2000SEP08	g2035952	2121	2368
11	LI:072072.1:2000SEP08	g1138172	321	623
11	LI:072072.1:2000SEP08	g3918602	323	704
11	LI:072072.1:2000SEP08	8118160H1	1	508
11	LI:072072.1:2000SEP08	7459387H1	32	573
11	LI:072072.1:2000SEP08	7594566H1	68	661
11	LI:072072.1:2000SEP08	g7276911	321	782
11	LI:072072.1:2000SEP08	71677001V1	1936	2597
11	LI:072072.1:2000SEP08	7952295H2	2109	2730
11	LI:072072.1:2000SEP08	5626631R8	2114	2637
11	LI:072072.1:2000SEP08	g1494233	2552	2812
11	LI:072072.1:2000SEP08	6388364F8	2493	2639
11	LI:072072.1:2000SEP08	71679256V1	2531	2885
11	LI:072072.1:2000SEP08	71681012V1	2537	3209
11	LI:072072.1:2000SEP08	g2167858	862	1363
11	LI:072072.1:2000SEP08	5968471H1	894	1444
11	LI:072072.1:2000SEP08	6922455H1	1101	1573
11	LI:072072.1:2000SEP08	6609645H1	2444	2858
11	LI:072072.1:2000SEP08	7380154H1	2457	2930
11	LI:072072.1:2000SEP08	7601926H1	1897	2483
11	LI:072072.1:2000SEP08	7111074H1	547	982
11	LI:072072.1:2000SEP08	2729623F6	687	1145
11	LI:072072.1:2000SEP08	2729623H1	687	942
11	LI:072072.1:2000SEP08	g1156558	751	1080
11	LI:072072.1:2000SEP08	3377873H1	1499	1687
11	LI:072072.1:2000SEP08	2133730F6	1673	2027
11	LI:072072.1:2000SEP08	2133730H1	1673	1943
11	LI:072072.1:2000SEP08	2531179F7	418	625
11	LI:072072.1:2000SEP08	6344340H1	420	695
11	LI:072072.1:2000SEP08	7645365J1	1948	2340
11	LI:072072.1:2000SEP08	2312253H1	1939	2191
11	LI:072072.1:2000SEP08	2741734F6	2023	2365
11	LI:072072.1:2000SEP08	7645365H1	2139	2769
11	LI:072072.1:2000SEP08	71679742V1	1117	1787
11	LI:072072.1:2000SEP08	g4703402	1139	1586
11	LI:072072.1:2000SEP08	6388364H1	2397	2637
11	LI:072072.1:2000SEP08	71680692V1	2348	2918
11	LI:072072.1:2000SEP08	6242978H1	2382	2637
11	LI:072072.1:2000SEP08	5626631H1	2114	2436

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
11	LI:072072.1:2000SEP08	5626631F8	2127	2654
11	LI:072072.1:2000SEP08	6583287H1	2109	2782
11	LI:072072.1:2000SEP08	6112531F8	2081	2704
11	LI:072072.1:2000SEP08	g994115	1292	1505
11	LI:072072.1:2000SEP08	3258260H1	1294	1582
11	LI:072072.1:2000SEP08	2733768H1	2023	2285
11	LI:072072.1:2000SEP08	6112531H1	2077	2343
11	LI:072072.1:2000SEP08	2732620H1	2023	2282
11	LI:072072.1:2000SEP08	7003046H1	1703	2297
11	LI:072072.1:2000SEP08	6449756H1	1869	2458
11	LI:072072.1:2000SEP08	2312253R6	1939	2368
11	LI:072072.1:2000SEP08	71677869V1	2672	3188
11	LI:072072.1:2000SEP08	6141277H1	2672	2915
11	LI:072072.1:2000SEP08	g5765585	2672	2904
11	LI:072072.1:2000SEP08	71678512V1	2672	3305
11	LI:072072.1:2000SEP08	5722947H1	2672	2858
11	LI:072072.1:2000SEP08	71678780V1	2672	2886
11	LI:072072.1:2000SEP08	71679319V1	2672	2886
11	LI:072072.1:2000SEP08	3216015F6	2784	2910
11	LI:072072.1:2000SEP08	3216015H1	2784	2858
12	LI:148565.4:2000SEP08	6269049H1	1	572
12	LI:148565.4:2000SEP08	71608134V1	1045	1517
12	LI:148565.4:2000SEP08	4108727H1	29	275
12	LI:148565.4:2000SEP08	71607971V1	80	683
12	LI:148565.4:2000SEP08	5912118F8	80	616
12	LI:148565.4:2000SEP08	71608037V1	80	568
12	LI:148565.4:2000SEP08	71607529V1	80	650
12	LI:148565.4:2000SEP08	71603665V1	80	768
12	LI:148565.4:2000SEP08	71608875V1	80	797
12	LI:148565.4:2000SEP08	71605715V1	80	555
12	LI:148565.4:2000SEP08	5912118F6	82	437
12	LI:148565.4:2000SEP08	5912002F6	113	416
12	LI:148565.4:2000SEP08	5912002F8	114	624
12	LI:148565.4:2000SEP08	5912118H1	168	390
12	LI:148565.4:2000SEP08	5912002H1	168	364
12	LI:148565.4:2000SEP08	71608189V1	484	1108
12	LI:148565.4:2000SEP08	71609047V1	576	1224
12	LI:148565.4:2000SEP08	71604596V1	722	1472
12	LI:148565.4:2000SEP08	71605080V1	853	1535
12	LI:148565.4:2000SEP08	71604020V1	922	1534
12	LI:148565.4:2000SEP08	71605566V1	983	1535
12	LI:148565.4:2000SEP08	71603835V1	974	1532
12	LI:148565.4:2000SEP08	6269049T8	967	1345
12	LI:148565.4:2000SEP08	6269049F8	1	650
13	LI:368626.4:2000SEP08	g6710138	999	1340
13	LI:368626.4:2000SEP08	55031278H1	1	668
13	LI:368626.4:2000SEP08	55031280H1	1	672
13	LI:368626.4:2000SEP08	55031277H1	1	679
13	LI:368626.4:2000SEP08	55031279H1	1	668

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
13	LI:368626.4:2000SEP08	55031275H1	1	681
13	LI:368626.4:2000SEP08	55031274H1	1	524
13	LI:368626.4:2000SEP08	55031273H1	3	679
13	LI:368626.4:2000SEP08	6059113H1	143	694
13	LI:368626.4:2000SEP08	5096451F6	284	562
13	LI:368626.4:2000SEP08	2041326H1	324	584
13	LI:368626.4:2000SEP08	55031278T1	571	1225
13	LI:368626.4:2000SEP08	55031273T1	580	1165
13	LI:368626.4:2000SEP08	55031274T1	579	1218
13	LI:368626.4:2000SEP08	55031279T1	609	1157
13	LI:368626.4:2000SEP08	55031280T1	627	1113
13	LI:368626.4:2000SEP08	2621425T6	644	1301
13	LI:368626.4:2000SEP08	55031276T1	660	1093
13	LI:368626.4:2000SEP08	55031277T1	716	1108
13	LI:368626.4:2000SEP08	7256734H1	754	1357
13	LI:368626.4:2000SEP08	7256834H1	765	1340
13	LI:368626.4:2000SEP08	7256634H1	851	1340
14	LI:346123.1:2000SEP08	71639403V1	8	492
14	LI:346123.1:2000SEP08	4940021H1	8	150
14	LI:346123.1:2000SEP08	71635173V1	44	510
14	LI:346123.1:2000SEP08	71636893V1	44	509
14	LI:346123.1:2000SEP08	71636855V1	44	684
14	LI:346123.1:2000SEP08	71634240V1	44	710
14	LI:346123.1:2000SEP08	5909342F6	44	258
14	LI:346123.1:2000SEP08	4940021T9	592	1016
14	LI:346123.1:2000SEP08	71633087V1	662	911
14	LI:346123.1:2000SEP08	4142617H1	621	938
14	LI:346123.1:2000SEP08	71628325V1	705	890
14	LI:346123.1:2000SEP08	71636820V1	642	1065
14	LI:346123.1:2000SEP08	71634191V1	44	549
14	LI:346123.1:2000SEP08	71633913V1	44	488
14	LI:346123.1:2000SEP08	71634147V1	44	740
14	LI:346123.1:2000SEP08	71634091V1	8	546
14	LI:346123.1:2000SEP08	71635385V1	8	509
14	LI:346123.1:2000SEP08	71638406V1	8	439
14	LI:346123.1:2000SEP08	4940021F7	33	553
14	LI:346123.1:2000SEP08	71635657V1	8	589
14	LI:346123.1:2000SEP08	71637949V1	44	494
14	LI:346123.1:2000SEP08	71630147V1	44	145
14	LI:346123.1:2000SEP08	71635742V1	166	691
14	LI:346123.1:2000SEP08	71634782V1	44	687
14	LI:346123.1:2000SEP08	71635417V1	44	720
14	LI:346123.1:2000SEP08	71633942V1	44	624
14	LI:346123.1:2000SEP08	71634825V1	201	746
14	LI:346123.1:2000SEP08	71637654V1	283	907
14	LI:346123.1:2000SEP08	71637472V1	44	545
14	LI:346123.1:2000SEP08	71635361V1	44	621
14	LI:346123.1:2000SEP08	71636906V1	44	590
14	LI:346123.1:2000SEP08	5909342F8	44	565

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
14	LI:346123.1:2000SEP08	71635936V1	44	446
14	LI:346123.1:2000SEP08	71637681V1	44	435
14	LI:346123.1:2000SEP08	71634551V1	339	923
14	LI:346123.1:2000SEP08	71637225V1	587	1116
14	LI:346123.1:2000SEP08	71634188V1	1	588
14	LI:346123.1:2000SEP08	71634234V1	44	508
14	LI:346123.1:2000SEP08	71638668V1	44	509
14	LI:346123.1:2000SEP08	5909342H1	44	305
15	LI:335795.11:2000SEP08	5501337H1	1731	1984
15	LI:335795.11:2000SEP08	g1192132	2915	3230
15	LI:335795.11:2000SEP08	g820818	2950	3235
15	LI:335795.11:2000SEP08	3052621H1	2945	3231
15	LI:335795.11:2000SEP08	g4078441	2959	3414
15	LI:335795.11:2000SEP08	g1149247	3015	3228
15	LI:335795.11:2000SEP08	71031243V1	2436	3112
15	LI:335795.11:2000SEP08	4445762H1	2435	2685
15	LI:335795.11:2000SEP08	5760179H1	1532	1795
15	LI:335795.11:2000SEP08	6495832H1	1544	2166
15	LI:335795.11:2000SEP08	7113995H1	1880	2485
15	LI:335795.11:2000SEP08	5329068H1	2059	2319
15	LI:335795.11:2000SEP08	70911787V1	2592	3205
15	LI:335795.11:2000SEP08	g3240207	70	399
15	LI:335795.11:2000SEP08	6781168H1	1129	1543
15	LI:335795.11:2000SEP08	2290180R6	2819	3056
15	LI:335795.11:2000SEP08	5310325H1	2829	3096
15	LI:335795.11:2000SEP08	g1162453	3038	3227
15	LI:335795.11:2000SEP08	3368787H1	3072	3227
15	LI:335795.11:2000SEP08	g875549	3163	3231
15	LI:335795.11:2000SEP08	5672928F8	2586	3068
15	LI:335795.11:2000SEP08	6353715H1	2410	2701
15	LI:335795.11:2000SEP08	70911318V1	2428	2846
15	LI:335795.11:2000SEP08	4846483T6	2494	2966
15	LI:335795.11:2000SEP08	g1516482	2506	2923
15	LI:335795.11:2000SEP08	1623494T6	2480	2882
15	LI:335795.11:2000SEP08	6778208J1	378	1010
15	LI:335795.11:2000SEP08	g2985527	577	916
15	LI:335795.11:2000SEP08	70911560V1	2403	3101
15	LI:335795.11:2000SEP08	g6229165	99	397
15	LI:335795.11:2000SEP08	3282810F6	281	712
15	LI:335795.11:2000SEP08	3282810H1	281	530
15	LI:335795.11:2000SEP08	5691909H1	2324	2632
15	LI:335795.11:2000SEP08	70911658V1	2346	2848
15	LI:335795.11:2000SEP08	7116005H2	1880	2531
15	LI:335795.11:2000SEP08	g4969831	2395	2867
15	LI:335795.11:2000SEP08	3171081H1	2395	2680
15	LI:335795.11:2000SEP08	3171374H1	2396	2681
15	LI:335795.11:2000SEP08	1287842F1	2402	2974
15	LI:335795.11:2000SEP08	4743929H1	1270	1497
15	LI:335795.11:2000SEP08	5351789H1	1398	1613



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
15	LI:335795.11:2000SEP08	5883285H1	2798	3131
15	LI:335795.11:2000SEP08	2290180T6	2819	3190
15	LI:335795.11:2000SEP08	3246061H1	930	1185
15	LI:335795.11:2000SEP08	2245068H1	2385	2637
15	LI:335795.11:2000SEP08	3959886H2	1629	1906
15	LI:335795.11:2000SEP08	7114191H2	1880	2325
15	LI:335795.11:2000SEP08	2375912H1	1867	2130
15	LI:335795.11:2000SEP08	4639879H1	1849	2099
15	LI:335795.11:2000SEP08	2881484F6	1589	2066
15	LI:335795.11:2000SEP08	70914875V1	2716	3243
15	LI:335795.11:2000SEP08	g5744770	2542	2931
15	LI:335795.11:2000SEP08	5610413H1	2543	2747
15	LI:335795.11:2000SEP08	70913456V1	2577	3227
15	LI:335795.11:2000SEP08	g4137530	2541	2929
15	LI:335795.11:2000SEP08	g3052971	2773	3227
15	LI:335795.11:2000SEP08	g7156689	2794	3229
15	LI:335795.11:2000SEP08	5881091H1	2797	3082
15	LI:335795.11:2000SEP08	g6575589	2737	3226
15	LI:335795.11:2000SEP08	g2155531	2744	3230
15	LI:335795.11:2000SEP08	1368211H1	2192	2445
15	LI:335795.11:2000SEP08	1368211R1	2192	2671
15	LI:335795.11:2000SEP08	g1516483	2206	2651
15	LI:335795.11:2000SEP08	7113914H1	1880	2419
15	LI:335795.11:2000SEP08	7113460H1	1880	2359
15	LI:335795.11:2000SEP08	70913301V1	1867	2459
15	LI:335795.11:2000SEP08	6518284H1	1837	2391
15	LI:335795.11:2000SEP08	2375912F6	1867	2457
15	LI:335795.11:2000SEP08	71272229V1	1867	2494
15	LI:335795.11:2000SEP08	3187232H1	2453	2783
15	LI:335795.11:2000SEP08	5271263H1	2453	2715
15	LI:335795.11:2000SEP08	g3433280	2464	2923
15	LI:335795.11:2000SEP08	g1187857	2450	2606
15	LI:335795.11:2000SEP08	7613963H1	2698	3010
15	LI:335795.11:2000SEP08	5438870T9	2711	3320
15	LI:335795.11:2000SEP08	g2834848	2692	2923
15	LI:335795.11:2000SEP08	432992H1	2692	2911
15	LI:335795.11:2000SEP08	2290180H1	2831	3110
15	LI:335795.11:2000SEP08	g3098814	2842	3227
15	LI:335795.11:2000SEP08	g4876651	2884	2950
15	LI:335795.11:2000SEP08	796575H1	2891	3176
15	LI:335795.11:2000SEP08	g5529523	2908	3227
15	LI:335795.11:2000SEP08	70913232V1	2661	3226
15	LI:335795.11:2000SEP08	5468292H1	2665	2897
15	LI:335795.11:2000SEP08	2347807H1	2675	2931
15	LI:335795.11:2000SEP08	71666722V1	2650	2760
15	LI:335795.11:2000SEP08	2375912T6	2661	3188
15	LI:335795.11:2000SEP08	1623494F6	2160	2620
15	LI:335795.11:2000SEP08	1623494H1	2160	2392
15	LI:335795.11:2000SEP08	2972467H2	2184	2495

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
15	LI:335795.11:2000SEP08	3881684H1	2190	2484
15	LI:335795.11:2000SEP08	1287842H1	2402	2669
15	LI:335795.11:2000SEP08	6944693H1	1728	2354
15	LI:335795.11:2000SEP08	g2018726	2008	2383
15	LI:335795.11:2000SEP08	6888836J1	2219	2850
15	LI:335795.11:2000SEP08	2704991H1	2230	2526
15	LI:335795.11:2000SEP08	7646995H1	1604	1746
15	LI:335795.11:2000SEP08	6888836H1	1644	2135
15	LI:335795.11:2000SEP08	5189887H1	2148	2390
15	LI:335795.11:2000SEP08	5831568H2	383	627
15	LI:335795.11:2000SEP08	6379449H1	432	718
15	LI:335795.11:2000SEP08	7112844H2	1880	2455
15	LI:335795.11:2000SEP08	5376792H1	2358	2576
15	LI:335795.11:2000SEP08	g6300247	1	198
15	LI:335795.11:2000SEP08	g6302837	1	200
15	LI:335795.11:2000SEP08	4959826T8	14	284
15	LI:335795.11:2000SEP08	71271067V1	2055	2685
15	LI:335795.11:2000SEP08	3526296H1	2437	2727
15	LI:335795.11:2000SEP08	5267964H1	2437	2708
15	LI:335795.11:2000SEP08	1491592H1	2443	2648
15	LI:335795.11:2000SEP08	g1165370	2450	2813
15	LI:335795.11:2000SEP08	2918889T6	2388	2956
15	LI:335795.11:2000SEP08	5374261T6	2354	2891
15	LI:335795.11:2000SEP08	7114829H1	1880	2228
15	LI:335795.11:2000SEP08	70911039V1	1867	2479
15	LI:335795.11:2000SEP08	71271501V1	1867	2500
15	LI:335795.11:2000SEP08	70913941V1	1866	2460
15	LI:335795.11:2000SEP08	4959826F8	14	404
15	LI:335795.11:2000SEP08	g6229013	62	397
15	LI:335795.11:2000SEP08	5681962H1	2232	2513
15	LI:335795.11:2000SEP08	3089367H1	2259	2541
15	LI:335795.11:2000SEP08	1787665H1	2256	2311
15	LI:335795.11:2000SEP08	71271124V1	2259	2848
15	LI:335795.11:2000SEP08	7032667H1	2271	2784
15	LI:335795.11:2000SEP08	7345758H1	2275	2886
15	LI:335795.11:2000SEP08	3451289T6	2282	2865
15	LI:335795.11:2000SEP08	3012053T6	2294	2887
15	LI:335795.11:2000SEP08	3761546H1	2295	2636
15	LI:335795.11:2000SEP08	2881484H1	1589	1867
15	LI:335795.11:2000SEP08	7134302H1	1903	2436
15	LI:335795.11:2000SEP08	g5232793	1960	2293
15	LI:335795.11:2000SEP08	625726H1	1982	2277
15	LI:335795.11:2000SEP08	70912155V1	1997	2607
15	LI:335795.11:2000SEP08	3468658H1	1998	2298
15	LI:335795.11:2000SEP08	7115807H2	1881	2547
15	LI:335795.11:2000SEP08	7113392H1	1885	2253
15	LI:335795.11:2000SEP08	71271025V1	1890	2400
15	LI:335795.11:2000SEP08	5593330H1	2082	2287
15	LI:335795.11:2000SEP08	5268570H1	2145	2423

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
15	LI:335795.11:2000SEP08	3801457H1	1112	1357
15	LI:335795.11:2000SEP08	60111227B1	2597	3200
15	LI:335795.11:2000SEP08	2881484T6	2597	3202
15	LI:335795.11:2000SEP08	g2141943	2648	3162
15	LI:335795.11:2000SEP08	71667206V1	2650	2762
15	LI:335795.11:2000SEP08	71667114V1	2650	2817
15	LI:335795.11:2000SEP08	6046416H1	1107	1627
15	LI:335795.11:2000SEP08	6046416J1	1183	1735
15	LI:335795.11:2000SEP08	3012053F6	1214	1615
15	LI:335795.11:2000SEP08	5376784H1	1231	1458
16	LI:246023.2:2000SEP08	71081207V1	433	1065
16	LI:246023.2:2000SEP08	8113269H1	442	1036
16	LI:246023.2:2000SEP08	7435763H1	446	1007
16	LI:246023.2:2000SEP08	71253312V1	441	1089
16	LI:246023.2:2000SEP08	71252981V1	293	932
16	LI:246023.2:2000SEP08	6567190H1	317	846
16	LI:246023.2:2000SEP08	859574H1	317	552
16	LI:246023.2:2000SEP08	3761187H1	353	561
16	LI:246023.2:2000SEP08	4863263H1	428	691
16	LI:246023.2:2000SEP08	71253372V1	435	1109
16	LI:246023.2:2000SEP08	6610019H1	49	566
16	LI:246023.2:2000SEP08	6022017H1	731	1044
16	LI:246023.2:2000SEP08	3695673H1	760	1042
16	LI:246023.2:2000SEP08	3520128H1	770	1040
16	LI:246023.2:2000SEP08	909243H1	793	896
16	LI:246023.2:2000SEP08	2627223H1	926	1157
16	LI:246023.2:2000SEP08	2728422H1	10	254
16	LI:246023.2:2000SEP08	2662990H1	11	238
16	LI:246023.2:2000SEP08	7716637H1	665	1281
16	LI:246023.2:2000SEP08	5660534F8	268	702
16	LI:246023.2:2000SEP08	71081365V1	285	894
16	LI:246023.2:2000SEP08	2663466H1	11	256
16	LI:246023.2:2000SEP08	4239284H1	12	289
16	LI:246023.2:2000SEP08	7729208H1	510	990
16	LI:246023.2:2000SEP08	71082814V1	539	1147
16	LI:246023.2:2000SEP08	7674440H2	542	1026
16	LI:246023.2:2000SEP08	7715837J1	589	1233
16	LI:246023.2:2000SEP08	7716637J1	589	1216
16	LI:246023.2:2000SEP08	7325365H1	602	1173
16	LI:246023.2:2000SEP08	71252939V1	606	1185
16	LI:246023.2:2000SEP08	3077475F7	616	869
16	LI:246023.2:2000SEP08	7726911H1	624	1124
16	LI:246023.2:2000SEP08	71252962V1	642	1212
16	LI:246023.2:2000SEP08	5896708H1	645	970
16	LI:246023.2:2000SEP08	3044820H1	40	314
16	LI:246023.2:2000SEP08	71083664V1	47	658
16	LI:246023.2:2000SEP08	7579683H1	48	496
16	LI:246023.2:2000SEP08	3225085H1	43	314
16	LI:246023.2:2000SEP08	7217489H1	49	507

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
16	LI:246023.2:2000SEP08	4982031H1	50	203
16	LI:246023.2:2000SEP08	2719787H1	2	233
16	LI:246023.2:2000SEP08	7589302H2	7	574
16	LI:246023.2:2000SEP08	2728422F6	10	513
16	LI:246023.2:2000SEP08	6077914F8	1	626
16	LI:246023.2:2000SEP08	3634404H1	1	309
16	LI:246023.2:2000SEP08	6077914H1	1	326
16	LI:246023.2:2000SEP08	7003247H1	49	605
16	LI:246023.2:2000SEP08	3119527H1	45	140
16	LI:246023.2:2000SEP08	880437H1	50	193
16	LI:246023.2:2000SEP08	4158283H1	67	311
16	LI:246023.2:2000SEP08	7219095H1	71	623
16	LI:246023.2:2000SEP08	5660534H1	199	439
16	LI:246023.2:2000SEP08	7641994J1	211	468
16	LI:246023.2:2000SEP08	7641994H1	211	497
16	LI:246023.2:2000SEP08	3348779H1	216	473
16	LI:246023.2:2000SEP08	71089987V1	225	508
16	LI:246023.2:2000SEP08	6609320H1	236	740
16	LI:246023.2:2000SEP08	7262974H1	49	549
16	LI:246023.2:2000SEP08	7169252H1	47	597
16	LI:246023.2:2000SEP08	3448653H1	48	199
16	LI:246023.2:2000SEP08	3147889H1	50	276
16	LI:246023.2:2000SEP08	4448069H1	50	315
16	LI:246023.2:2000SEP08	880437R1	50	637
17	LG:1100661.1:2000SEP08	6790680H1	1	294
17	LG:1100661.1:2000SEP08	6791384H1	1	293
17	LG:1100661.1:2000SEP08	6798074F8	3	459
17	LG:1100661.1:2000SEP08	6798074H1	3	308
17	LG:1100661.1:2000SEP08	6792683T8	14	407
17	LG:1100661.1:2000SEP08	6792683F8	7	457
17	LG:1100661.1:2000SEP08	6795817F8	9	514
17	LG:1100661.1:2000SEP08	6792683H1	9	348
17	LG:1100661.1:2000SEP08	6795817H1	15	311
17	LG:1100661.1:2000SEP08	6795817T8	249	414
17	LG:1100661.1:2000SEP08	6790621H1	263	466
17	LG:1100661.1:2000SEP08	6794755F8	263	466
17	LG:1100661.1:2000SEP08	6794755H1	263	466
18	LG:475856.1:2000SEP08	7761183H1	1	614
18	LG:475856.1:2000SEP08	5779813H1	373	493
18	LG:475856.1:2000SEP08	5779913H1	374	493
18	LG:475856.1:2000SEP08	5779813T6	373	883
19	LG:1015343.1:2000SEP08	6798647T8	1	541
19	LG:1015343.1:2000SEP08	6798647H1	1	434
19	LG:1015343.1:2000SEP08	6798647F8	1	615
20	LG:1400575.1:2000SEP08	5494669R6	596	908
20	LG:1400575.1:2000SEP08	4645588F9	1	571
20	LG:1400575.1:2000SEP08	7273781H1	8	564
20	LG:1400575.1:2000SEP08	2383223F6	11	483
20	LG:1400575.1:2000SEP08	7467241H1	24	579

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
20	LG:1400575.1:2000SEP08	4289889F8	81	689
20	LG:1400575.1:2000SEP08	4289889H1	81	333
20	LG:1400575.1:2000SEP08	7422342T1	177	630
20	LG:1400575.1:2000SEP08	6823362J1	340	964
20	LG:1400575.1:2000SEP08	6823362H1	339	911
21	LG:1080545.1:2000SEP08	1546821R6	489	625
21	LG:1080545.1:2000SEP08	2992129T6	184	633
21	LG:1080545.1:2000SEP08	2992129H1	34	317
21	LG:1080545.1:2000SEP08	g3756177	429	632
21	LG:1080545.1:2000SEP08	g6568887	360	632
21	LG:1080545.1:2000SEP08	5283246F7	285	581
21	LG:1080545.1:2000SEP08	5995749T8	222	651
21	LG:1080545.1:2000SEP08	6571393T8	528	654
21	LG:1080545.1:2000SEP08	5435937T8	462	628
21	LG:1080545.1:2000SEP08	5019177T6	198	633
21	LG:1080545.1:2000SEP08	1546821T6	477	633
21	LG:1080545.1:2000SEP08	6571393F8	537	632
21	LG:1080545.1:2000SEP08	7759319J1	1	518
21	LG:1080545.1:2000SEP08	3803988H1	72	376
21	LG:1080545.1:2000SEP08	1546821H1	489	672
21	LG:1080545.1:2000SEP08	6412272T8	380	658
21	LG:1080545.1:2000SEP08	2806609T6	272	649
21	LG:1080545.1:2000SEP08	6364173H1	247	411
21	LG:1080545.1:2000SEP08	7759319H1	2	367
21	LG:1080545.1:2000SEP08	2992129F6	34	485
21	LG:1080545.1:2000SEP08	7589990H1	2	351
21	LG:1080545.1:2000SEP08	6571393H1	537	632
22	LG:213947.1:2000SEP08	g1194144	50	378
22	LG:213947.1:2000SEP08	g1059990	81	362
22	LG:213947.1:2000SEP08	3605933H1	1	125
23	LI:720641.1:2000SEP08	6561042H1	1	543
23	LI:720641.1:2000SEP08	6561042F8	1	683
23	LI:720641.1:2000SEP08	6561042T8	380	1033
24	LI:1023894.1:2000SEP08	6796705F8	1	608
24	LI:1023894.1:2000SEP08	6796705H1	1	459
24	LI:1023894.1:2000SEP08	6796705T8	454	963
25	LI:734904.1:2000SEP08	7407034H1	974	1519
25	LI:734904.1:2000SEP08	5091005R8	986	1214
25	LI:734904.1:2000SEP08	71175220V1	292	931
25	LI:734904.1:2000SEP08	55001375J1	1186	1799
25	LI:734904.1:2000SEP08	55001376H1	1186	1794
25	LI:734904.1:2000SEP08	55001374H1	1187	1788
25	LI:734904.1:2000SEP08	55001372H1	1186	1454
25	LI:734904.1:2000SEP08	55001371J1	1186	1799
25	LI:734904.1:2000SEP08	55001373H1	1186	1789
25	LI:734904.1:2000SEP08	55001376J1	1186	1799
25	LI:734904.1:2000SEP08	71498568V1	1144	1410
25	LI:734904.1:2000SEP08	71469766V1	1145	1402
25	LI:734904.1:2000SEP08	55001375H1	1186	1735

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
25	LI:734904.1:2000SEP08	5080305H1	460	620
25	LI:734904.1:2000SEP08	55001377H1	1186	1785
25	LI:734904.1:2000SEP08	55001371H1	1186	1810
25	LI:734904.1:2000SEP08	g4069540	324	760
25	LI:734904.1:2000SEP08	5954166H1	334	675
25	LI:734904.1:2000SEP08	5954266H1	447	675
25	LI:734904.1:2000SEP08	55001372J1	1186	1799
25	LI:734904.1:2000SEP08	55001373J1	1186	1799
25	LI:734904.1:2000SEP08	55001380H1	1186	1794
25	LI:734904.1:2000SEP08	55001380J1	1186	1798
25	LI:734904.1:2000SEP08	55001374J1	1186	1789
25	LI:734904.1:2000SEP08	5155483H1	444	713
25	LI:734904.1:2000SEP08	71470137V1	1145	1668
25	LI:734904.1:2000SEP08	71468740V1	1145	1670
25	LI:734904.1:2000SEP08	71470749V1	1145	1612
25	LI:734904.1:2000SEP08	71471307V1	1145	1589
25	LI:734904.1:2000SEP08	71471777V1	1145	1553
25	LI:734904.1:2000SEP08	71471014V1	1145	1498
25	LI:734904.1:2000SEP08	71528917V1	1145	1473
25	LI:734904.1:2000SEP08	55009941J1	1077	1265
25	LI:734904.1:2000SEP08	7726872H1	1166	1614
25	LI:734904.1:2000SEP08	5090822F6	1168	1618
25	LI:734904.1:2000SEP08	71469065V1	1143	1712
25	LI:734904.1:2000SEP08	71472073V1	1144	1661
25	LI:734904.1:2000SEP08	8100283H1	969	1556
25	LI:734904.1:2000SEP08	2685561H1	8	109
25	LI:734904.1:2000SEP08	1921627H1	91	347
25	LI:734904.1:2000SEP08	2767577H1	1	184
25	LI:734904.1:2000SEP08	g1921766	389	812
25	LI:734904.1:2000SEP08	55001379H1	1186	1546
25	LI:734904.1:2000SEP08	5734758F6	292	886
25	LI:734904.1:2000SEP08	5734758H1	292	571
25	LI:734904.1:2000SEP08	g3214644	319	743
25	LI:734904.1:2000SEP08	5807132H1	190	426
25	LI:734904.1:2000SEP08	2554815H1	205	445
25	LI:734904.1:2000SEP08	55001378H1	1187	1759
25	LI:734904.1:2000SEP08	866814H1	98	406
25	LI:734904.1:2000SEP08	g6836844	1599	1975
25	LI:734904.1:2000SEP08	3483164H1	1645	1895
25	LI:734904.1:2000SEP08	60211015U1	501	1050
25	LI:734904.1:2000SEP08	g7157221	1581	1994
25	LI:734904.1:2000SEP08	5136773H2	1534	1802
25	LI:734904.1:2000SEP08	g7237043	1536	1990
25	LI:734904.1:2000SEP08	71468287V1	1560	2008
25	LI:734904.1:2000SEP08	7618825J1	1565	1947
25	LI:734904.1:2000SEP08	71470403V1	1145	1771
25	LI:734904.1:2000SEP08	8103182H1	951	1601
25	LI:734904.1:2000SEP08	7736288J1	496	1110
25	LI:734904.1:2000SEP08	2914008H1	641	787

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
25	LI:734904.1:2000SEP08	7729292J1	898	1429
25	LI:734904.1:2000SEP08	7120533F8	920	1075
25	LI:734904.1:2000SEP08	7702456H1	909	1233
25	LI:734904.1:2000SEP08	55002849H1	921	1135
25	LI:734904.1:2000SEP08	7399819H1	916	1422
25	LI:734904.1:2000SEP08	g4070593	320	773
25	LI:734904.1:2000SEP08	g2318943	322	619
25	LI:734904.1:2000SEP08	7729273J1	922	1425
25	LI:734904.1:2000SEP08	7932828H1	928	1081
25	LI:734904.1:2000SEP08	7120533H1	964	1102
25	LI:734904.1:2000SEP08	8104231H1	921	984
25	LI:734904.1:2000SEP08	7610629H1	955	1405
25	LI:734904.1:2000SEP08	7412449H1	1028	1299
25	LI:734904.1:2000SEP08	55009941H1	1056	1242
25	LI:734904.1:2000SEP08	8088518H1	1189	1730
25	LI:734904.1:2000SEP08	71618335V1	1207	1686
25	LI:734904.1:2000SEP08	5090822H1	1327	1618
25	LI:734904.1:2000SEP08	5091005H1	1342	1618
25	LI:734904.1:2000SEP08	7427190H1	1427	1818
25	LI:734904.1:2000SEP08	8105596H1	1508	1975
25	LI:734904.1:2000SEP08	8105596J1	1509	1984
25	LI:734904.1:2000SEP08	5613660H1	1517	1794
25	LI:734904.1:2000SEP08	g5769489	390	598
25	LI:734904.1:2000SEP08	5090425R6	987	1315
25	LI:734904.1:2000SEP08	7412451H1	1028	1300
25	LI:734904.1:2000SEP08	5953866H1	525	675
25	LI:734904.1:2000SEP08	3978021H1	558	727
26	LI:1178118.1:2000SEP08	70888593V1	1170	1676
26	LI:1178118.1:2000SEP08	70887574V1	1981	2421
26	LI:1178118.1:2000SEP08	70885508V1	1982	2459
26	LI:1178118.1:2000SEP08	70857604V1	1580	2173
26	LI:1178118.1:2000SEP08	70886023V1	897	1470
26	LI:1178118.1:2000SEP08	70886312V1	1535	2209
26	LI:1178118.1:2000SEP08	70887783V1	897	1464
26	LI:1178118.1:2000SEP08	70886381V1	897	1416
26	LI:1178118.1:2000SEP08	70885783V1	1981	2449
26	LI:1178118.1:2000SEP08	70887187V1	1981	2348
26	LI:1178118.1:2000SEP08	70863909V1	1436	1679
26	LI:1178118.1:2000SEP08	70887650V1	1981	2313
26	LI:1178118.1:2000SEP08	71226092V1	1279	1679
26	LI:1178118.1:2000SEP08	71225492V1	1408	1679
26	LI:1178118.1:2000SEP08	70886045V1	2304	2459
26	LI:1178118.1:2000SEP08	70855682V1	1981	2384
26	LI:1178118.1:2000SEP08	70856931V1	958	1596
26	LI:1178118.1:2000SEP08	70885518V1	897	1454
26	LI:1178118.1:2000SEP08	70885501V1	928	1554
26	LI:1178118.1:2000SEP08	70886925V1	1981	2186
26	LI:1178118.1:2000SEP08	70885179V1	1441	1676
26	LI:1178118.1:2000SEP08	70888682V1	1234	1679

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
26	LI:1178118.1:2000SEP08	70885893V1	1270	1460
26	LI:1178118.1:2000SEP08	70886871V1	1273	1679
26	LI:1178118.1:2000SEP08	71226596V1	1592	1679
26	LI:1178118.1:2000SEP08	71225755V1	2203	2459
26	LI:1178118.1:2000SEP08	70856256V1	1294	1679
26	LI:1178118.1:2000SEP08	70885056V1	1441	1676
26	LI:1178118.1:2000SEP08	70856069V1	1071	1519
26	LI:1178118.1:2000SEP08	70885472V1	1993	2470
26	LI:1178118.1:2000SEP08	70858123V1	1489	1679
26	LI:1178118.1:2000SEP08	70888070V1	897	1476
26	LI:1178118.1:2000SEP08	70885504V1	2008	2459
26	LI:1178118.1:2000SEP08	70856838V1	1480	2130
26	LI:1178118.1:2000SEP08	3918792H1	1981	2111
26	LI:1178118.1:2000SEP08	4246988H1	1982	2116
26	LI:1178118.1:2000SEP08	70856171V1	1981	2282
26	LI:1178118.1:2000SEP08	g1691161	1982	2342
26	LI:1178118.1:2000SEP08	2878289T6	1984	2435
26	LI:1178118.1:2000SEP08	5768870H1	1993	2459
26	LI:1178118.1:2000SEP08	70176041V1	2300	2487
26	LI:1178118.1:2000SEP08	g2079789	1995	2471
26	LI:1178118.1:2000SEP08	612729R6	2007	2297
26	LI:1178118.1:2000SEP08	612729H1	2007	2190
26	LI:1178118.1:2000SEP08	g2180032	2018	2409
26	LI:1178118.1:2000SEP08	70855176V1	2043	2470
26	LI:1178118.1:2000SEP08	g2566462	2091	2404
26	LI:1178118.1:2000SEP08	g4175126	2108	2483
26	LI:1178118.1:2000SEP08	g3433262	2108	2486
26	LI:1178118.1:2000SEP08	7737658J1	2142	2616
26	LI:1178118.1:2000SEP08	g832488	2155	2478
26	LI:1178118.1:2000SEP08	g1018319	2178	2472
26	LI:1178118.1:2000SEP08	g2594405	2216	2459
26	LI:1178118.1:2000SEP08	70856350V1	1579	2063
26	LI:1178118.1:2000SEP08	70856794V1	1547	2160
26	LI:1178118.1:2000SEP08	70858103V1	1358	1679
26	LI:1178118.1:2000SEP08	70856126V1	1457	1679
26	LI:1178118.1:2000SEP08	70885049V1	1449	1679
26	LI:1178118.1:2000SEP08	70855594V1	1371	1679
26	LI:1178118.1:2000SEP08	70887152V1	2015	2396
26	LI:1178118.1:2000SEP08	3918441H1	1981	2111
26	LI:1178118.1:2000SEP08	5508142R6	1981	2185
26	LI:1178118.1:2000SEP08	1242730H1	1981	2033
26	LI:1178118.1:2000SEP08	70818712V1	1378	1679
26	LI:1178118.1:2000SEP08	70886942V1	1204	1656
26	LI:1178118.1:2000SEP08	6880021F8	90	724
26	LI:1178118.1:2000SEP08	3493676F6	97	456
26	LI:1178118.1:2000SEP08	3493676H1	98	348
26	LI:1178118.1:2000SEP08	7155878H1	355	875
26	LI:1178118.1:2000SEP08	6176589H1	573	842
26	LI:1178118.1:2000SEP08	6880021J1	693	1325



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
26	U:1178118.1:2000SEP08	70857959V1	897	1500
26	U:1178118.1:2000SEP08	70885593V1	907	1552
26	U:1178118.1:2000SEP08	70886693V1	1066	1725
26	U:1178118.1:2000SEP08	70171420V1	1078	1402
26	U:1178118.1:2000SEP08	7155878J1	1208	1678
26	U:1178118.1:2000SEP08	70888002V1	1263	1683
26	U:1178118.1:2000SEP08	g766087	1	227
26	U:1178118.1:2000SEP08	70173885V1	1334	1685
26	U:1178118.1:2000SEP08	70173739V1	1374	1679
26	U:1178118.1:2000SEP08	70175935V1	1373	1715
26	U:1178118.1:2000SEP08	70887033V1	1403	1683
26	U:1178118.1:2000SEP08	6880021H2	90	567
26	U:1178118.1:2000SEP08	70886981V1	1414	1752
26	U:1178118.1:2000SEP08	70884826V1	1411	1788
26	U:1178118.1:2000SEP08	70171649V1	1429	1610
26	U:1178118.1:2000SEP08	71224949V1	1463	1679
26	U:1178118.1:2000SEP08	5508142F6	1456	1679
26	U:1178118.1:2000SEP08	5508142H1	1456	1657
26	U:1178118.1:2000SEP08	70887821V1	1450	1679
26	U:1178118.1:2000SEP08	71226480V1	1460	1725
26	U:1178118.1:2000SEP08	70886202V1	1544	2169
26	U:1178118.1:2000SEP08	70171399V1	1554	1676
26	U:1178118.1:2000SEP08	70172301V1	1554	1679
26	U:1178118.1:2000SEP08	g1018720	1567	1679
26	U:1178118.1:2000SEP08	70173180V1	1568	2094
26	U:1178118.1:2000SEP08	70886686V1	1574	2072
26	U:1178118.1:2000SEP08	3297526H1	1579	1676
26	U:1178118.1:2000SEP08	70171645V1	1604	2084
26	U:1178118.1:2000SEP08	2870408T6	1938	2428
26	U:1178118.1:2000SEP08	70171082V1	1955	2350
26	U:1178118.1:2000SEP08	g2079948	1962	2452
26	U:1178118.1:2000SEP08	70172284V1	1981	2065
26	U:1178118.1:2000SEP08	70175087V1	1981	2230
26	U:1178118.1:2000SEP08	70175279V1	1981	2230
26	U:1178118.1:2000SEP08	71225063V1	1981	2426
26	U:1178118.1:2000SEP08	70855560V1	1483	1679
26	U:1178118.1:2000SEP08	71226627V1	1305	1679
26	U:1178118.1:2000SEP08	7278205H1	462	949
26	U:1178118.1:2000SEP08	70888348V1	1981	2340
27	U:213947.1:2000SEP08	g1194144	1	329
27	U:213947.1:2000SEP08	g1059990	32	313
27	U:213947.1:2000SEP08	3605933H1	1	76
28	LG:407304.1:2000SEP08	6756208H1	1	234
28	LG:407304.1:2000SEP08	1393067T6	1	459
28	LG:407304.1:2000SEP08	6756208J1	1	661
28	LG:407304.1:2000SEP08	1393067F6	5	171
28	LG:407304.1:2000SEP08	1393067H1	4	152
28	LG:407304.1:2000SEP08	4881236H1	26	101
28	LG:407304.1:2000SEP08	5014037H1	50	291

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
28	LG:407304.1:2000SEP08	2179865H1	75	346
28	LG:407304.1:2000SEP08	4970021H1	162	432
28	LG:407304.1:2000SEP08	4706031T6	169	610
28	LG:407304.1:2000SEP08	7309340H1	199	496
28	LG:407304.1:2000SEP08	g5904385	232	695
28	LG:407304.1:2000SEP08	2669002H1	249	504
28	LG:407304.1:2000SEP08	g3307401	344	692
28	LG:407304.1:2000SEP08	7066680H1	507	618
28	LG:407304.1:2000SEP08	5423107H1	523	578
28	LG:407304.1:2000SEP08	5067631H1	524	816
28	LG:407304.1:2000SEP08	5422307H1	523	578
28	LG:407304.1:2000SEP08	5067631F6	524	752
28	LG:407304.1:2000SEP08	6534646H1	543	764
28	LG:407304.1:2000SEP08	6914529J1	590	661
28	LG:407304.1:2000SEP08	6914529H1	590	661
28	LG:407304.1:2000SEP08	440475H1	716	845
28	LG:407304.1:2000SEP08	g1894072	758	1079
29	LG:337358.1:2000SEP08	2244837H1	2959	3211
29	LG:337358.1:2000SEP08	6019194H1	2844	3383
29	LG:337358.1:2000SEP08	6019119H1	2845	3350
29	LG:337358.1:2000SEP08	4089922T6	2917	3177
29	LG:337358.1:2000SEP08	4090015T6	2950	3177
29	LG:337358.1:2000SEP08	1723941H1	2858	3068
29	LG:337358.1:2000SEP08	676035T6	2952	3362
29	LG:337358.1:2000SEP08	1723941F6	2858	3272
29	LG:337358.1:2000SEP08	4255978H1	1384	1624
29	LG:337358.1:2000SEP08	3447804H2	1426	1637
29	LG:337358.1:2000SEP08	7077206H1	1	342
29	LG:337358.1:2000SEP08	7582983H1	59	506
29	LG:337358.1:2000SEP08	7228466H1	260	871
29	LG:337358.1:2000SEP08	6698505H1	520	993
29	LG:337358.1:2000SEP08	6698505F8	520	937
29	LG:337358.1:2000SEP08	7125751F8	543	1072
29	LG:337358.1:2000SEP08	7125751H1	543	1019
29	LG:337358.1:2000SEP08	7178269H1	616	1097
29	LG:337358.1:2000SEP08	6054168H1	648	1198
29	LG:337358.1:2000SEP08	7074882H1	858	1407
29	LG:337358.1:2000SEP08	5386919H1	1167	1274
29	LG:337358.1:2000SEP08	7594269H1	1264	1853
29	LG:337358.1:2000SEP08	5680326H1	1334	1595
29	LG:337358.1:2000SEP08	2228689H1	1493	1738
29	LG:337358.1:2000SEP08	7180676H1	1499	2038
29	LG:337358.1:2000SEP08	5404677H1	1547	1716
29	LG:337358.1:2000SEP08	4091424H1	1576	1825
29	LG:337358.1:2000SEP08	5999641H1	1601	2107
29	LG:337358.1:2000SEP08	4090015F6	1672	2208
29	LG:337358.1:2000SEP08	4090015H1	1672	1939
29	LG:337358.1:2000SEP08	4837487H1	1702	1992
29	LG:337358.1:2000SEP08	5388090H1	1734	2034

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
29	LG:337358.1:2000SEP08	5662028H1	1741	1940
29	LG:337358.1:2000SEP08	4091549H1	1753	2019
29	LG:337358.1:2000SEP08	5664250H1	1783	2062
29	LG:337358.1:2000SEP08	3274759H1	1856	2105
29	LG:337358.1:2000SEP08	7074659H1	1918	2416
29	LG:337358.1:2000SEP08	676035H1	1940	2204
29	LG:337358.1:2000SEP08	676035R6	1940	2238
29	LG:337358.1:2000SEP08	7678201J1	1952	2366
29	LG:337358.1:2000SEP08	3969273H1	1980	2253
29	LG:337358.1:2000SEP08	g826435	2014	2316
29	LG:337358.1:2000SEP08	4091025H1	2078	2357
29	LG:337358.1:2000SEP08	2822209H1	2127	2432
29	LG:337358.1:2000SEP08	4785179H1	2197	2465
29	LG:337358.1:2000SEP08	5107431H1	2206	2276
29	LG:337358.1:2000SEP08	1818504F6	2254	2635
29	LG:337358.1:2000SEP08	1818504H1	2254	2533
29	LG:337358.1:2000SEP08	7119412H1	2257	2455
29	LG:337358.1:2000SEP08	6505231H1	2258	2465
29	LG:337358.1:2000SEP08	4129969H2	2336	2582
29	LG:337358.1:2000SEP08	5026554H1	2355	2602
29	LG:337358.1:2000SEP08	5998266H1	2385	2853
29	LG:337358.1:2000SEP08	776745H1	2411	2648
29	LG:337358.1:2000SEP08	776745R6	2411	2677
29	LG:337358.1:2000SEP08	775606R1	2411	2957
29	LG:337358.1:2000SEP08	775606H1	2411	2641
29	LG:337358.1:2000SEP08	5875073F6	2512	3100
29	LG:337358.1:2000SEP08	5875073H1	2513	2729
29	LG:337358.1:2000SEP08	6122560H1	2523	3099
29	LG:337358.1:2000SEP08	6129231H1	2523	2895
29	LG:337358.1:2000SEP08	2020216H1	2530	2635
29	LG:337358.1:2000SEP08	5692486H1	2571	2807
29	LG:337358.1:2000SEP08	5493633H1	2594	2882
29	LG:337358.1:2000SEP08	4093974H1	2631	2770
29	LG:337358.1:2000SEP08	4093981H1	2632	2821
29	LG:337358.1:2000SEP08	7237277H1	2632	2939
29	LG:337358.1:2000SEP08	4089922F6	2653	3048
29	LG:337358.1:2000SEP08	4089922H1	2653	2851
29	LG:337358.1:2000SEP08	3961245H2	2667	2865
29	LG:337358.1:2000SEP08	856602H1	2740	2913
29	LG:337358.1:2000SEP08	7129582H1	2754	3165
29	LG:337358.1:2000SEP08	3316652H1	2802	3114
29	LG:337358.1:2000SEP08	5758353H1	2819	3105
29	LG:337358.1:2000SEP08	4714095H1	2818	3045
29	LG:337358.1:2000SEP08	620821H1	2832	3106
29	LG:337358.1:2000SEP08	7342523H1	2996	3373
29	LG:337358.1:2000SEP08	5813971H1	2996	3204
29	LG:337358.1:2000SEP08	5823037H1	2996	3166
29	LG:337358.1:2000SEP08	5815923H1	2996	3186
29	LG:337358.1:2000SEP08	581695QH1	2997	3183

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
29	LG:337358.1:2000SEP08	7243161H2	3007	3194
29	LG:337358.1:2000SEP08	g4243969	3030	3410
29	LG:337358.1:2000SEP08	7250689H1	3029	3427
29	LG:337358.1:2000SEP08	4755254H1	3059	3329
29	LG:337358.1:2000SEP08	776745T6	3061	3361
29	LG:337358.1:2000SEP08	g826353	3134	3408
29	LG:337358.1:2000SEP08	5100636H1	3222	3377
29	LG:337358.1:2000SEP08	g1192066	3222	3417
29	LG:337358.1:2000SEP08	2571548H1	3247	3416
29	LG:337358.1:2000SEP08	5208905H1	3293	3410
30	LG:986090.1:2000SEP08	6269746F7	2	599
30	LG:986090.1:2000SEP08	6269746T7	459	831
30	LG:986090.1:2000SEP08	6271955F8	2	630
30	LG:986090.1:2000SEP08	6269746F8	1	654
30	LG:986090.1:2000SEP08	6271955H1	2	166
30	LG:986090.1:2000SEP08	6269746H1	1	475
31	LG:123250.1:2000SEP08	2640831H1	453	692
31	LG:123250.1:2000SEP08	2640844F6	453	1011
31	LG:123250.1:2000SEP08	2640844H1	453	684
31	LG:123250.1:2000SEP08	7647260J1	745	1364
31	LG:123250.1:2000SEP08	2640844T6	857	1055
31	LG:123250.1:2000SEP08	7413201H1	91	500
31	LG:123250.1:2000SEP08	1391077F6	50	450
31	LG:123250.1:2000SEP08	7647260H1	91	580
31	LG:123250.1:2000SEP08	1391077H1	50	281
31	LG:123250.1:2000SEP08	657586H1	37	281
31	LG:123250.1:2000SEP08	526846H1	30	278
31	LG:123250.1:2000SEP08	657586R7	38	558
31	LG:123250.1:2000SEP08	3336816H1	1	218
32	LG:1028774.2:2000SEP08	6754558J1	1066	1614
32	LG:1028774.2:2000SEP08	6768802J1	1068	1614
32	LG:1028774.2:2000SEP08	6753046J1	1146	1614
32	LG:1028774.2:2000SEP08	g2538197	1154	1621
32	LG:1028774.2:2000SEP08	g2270251	1323	1622
32	LG:1028774.2:2000SEP08	g3117302	1457	1622
32	LG:1028774.2:2000SEP08	g3412225	1281	1623
32	LG:1028774.2:2000SEP08	g3958404	1226	1604
32	LG:1028774.2:2000SEP08	6765755J1	1128	1613
32	LG:1028774.2:2000SEP08	6767391J1	1082	1613
32	LG:1028774.2:2000SEP08	6767336J1	1390	1614
32	LG:1028774.2:2000SEP08	6570749H1	982	1482
32	LG:1028774.2:2000SEP08	g6474933	1170	1571
32	LG:1028774.2:2000SEP08	g5768871	1170	1549
32	LG:1028774.2:2000SEP08	g6836462	1251	1571
32	LG:1028774.2:2000SEP08	g2910667	1191	1590
32	LG:1028774.2:2000SEP08	g5742470	1209	1597
32	LG:1028774.2:2000SEP08	g5590115	1168	1587
32	LG:1028774.2:2000SEP08	5645419F8	781	1107
32	LG:1028774.2:2000SEP08	4628319H1	818	1068

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
32	LG:1028774.2:2000SEP08	5645467H1	795	1036
32	LG:1028774.2:2000SEP08	3202082H1	956	1198
32	LG:1028774.2:2000SEP08	6319651H1	1011	1286
32	LG:1028774.2:2000SEP08	g2901003	1192	1363
32	LG:1028774.2:2000SEP08	6493972H1	45	549
32	LG:1028774.2:2000SEP08	7380087H1	155	655
32	LG:1028774.2:2000SEP08	6834528H1	57	616
32	LG:1028774.2:2000SEP08	268982H1	369	612
32	LG:1028774.2:2000SEP08	6783027H1	197	753
32	LG:1028774.2:2000SEP08	7260955H1	510	695
32	LG:1028774.2:2000SEP08	6131194H1	600	915
32	LG:1028774.2:2000SEP08	667680R6	69	352
32	LG:1028774.2:2000SEP08	667680H1	90	352
32	LG:1028774.2:2000SEP08	5195209H1	95	333
32	LG:1028774.2:2000SEP08	4623051H1	60	319
32	LG:1028774.2:2000SEP08	7656667J1	21	547
32	LG:1028774.2:2000SEP08	6753046H1	1	540
32	LG:1028774.2:2000SEP08	6245830H1	136	472
32	LG:1028774.2:2000SEP08	4699256H1	179	449
32	LG:1028774.2:2000SEP08	g5754453	1114	1623
32	LG:1028774.2:2000SEP08	g5365284	1132	1623
32	LG:1028774.2:2000SEP08	g3751680	1201	1623
32	LG:1028774.2:2000SEP08	g5887570	1126	1623
32	LG:1028774.2:2000SEP08	g5512642	1139	1623
33	LG:338927.6:2000SEP08	6810281H1	375	959
33	LG:338927.6:2000SEP08	g1195398	515	717
33	LG:338927.6:2000SEP08	g1192273	540	755
33	LG:338927.6:2000SEP08	7712077H1	624	1250
33	LG:338927.6:2000SEP08	g4189046	820	1087
33	LG:338927.6:2000SEP08	6804367H1	839	1254
33	LG:338927.6:2000SEP08	3341542H1	970	1210
33	LG:338927.6:2000SEP08	7712077J1	1	617
33	LG:338927.6:2000SEP08	6810281J1	220	846
34	LG:332944.2:2000SEP08	3288409H1	1602	1862
34	LG:332944.2:2000SEP08	3288409F6	1602	1857
34	LG:332944.2:2000SEP08	6582988H1	1677	2277
34	LG:332944.2:2000SEP08	6039693H1	1683	2235
34	LG:332944.2:2000SEP08	1952105H1	1691	1942
34	LG:332944.2:2000SEP08	7649251J2	1699	2318
34	LG:332944.2:2000SEP08	7054847H1	1703	2307
34	LG:332944.2:2000SEP08	6890168J1	1766	2401
34	LG:332944.2:2000SEP08	4216720H1	1804	2071
34	LG:332944.2:2000SEP08	094964H1	1872	2106
34	LG:332944.2:2000SEP08	2414057H1	1894	2140
34	LG:332944.2:2000SEP08	2414057R6	1894	2145
34	LG:332944.2:2000SEP08	1952625H1	1962	2209
34	LG:332944.2:2000SEP08	6912541H1	1967	2446
34	LG:332944.2:2000SEP08	g3419210	2360	2734
34	LG:332944.2:2000SEP08	4216326H1	1984	2252

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
34	LG:332944.2:2000SEP08	4216235H1	1985	2265
34	LG:332944.2:2000SEP08	6582988T1	2053	2674
34	LG:332944.2:2000SEP08	1948814H1	2063	2299
34	LG:332944.2:2000SEP08	2414057T6	2107	2696
34	LG:332944.2:2000SEP08	g1390588	2129	2436
34	LG:332944.2:2000SEP08	1759356H1	2156	2406
34	LG:332944.2:2000SEP08	626524H1	2163	2416
34	LG:332944.2:2000SEP08	620985H1	2388	2646
34	LG:332944.2:2000SEP08	2824851T6	2171	2701
34	LG:332944.2:2000SEP08	3490151T6	2206	2695
34	LG:332944.2:2000SEP08	2409354H1	2228	2460
34	LG:332944.2:2000SEP08	3288409T6	2240	2686
34	LG:332944.2:2000SEP08	g614376	2343	2732
34	LG:332944.2:2000SEP08	094296H1	694	989
34	LG:332944.2:2000SEP08	3494733H1	815	1094
34	LG:332944.2:2000SEP08	4775831F7	897	1402
34	LG:332944.2:2000SEP08	4775831H1	897	1133
34	LG:332944.2:2000SEP08	2824851F6	1067	1480
34	LG:332944.2:2000SEP08	2824851H1	1067	1295
34	LG:332944.2:2000SEP08	7185229H1	1095	1654
34	LG:332944.2:2000SEP08	7649969H2	1141	1773
34	LG:332944.2:2000SEP08	3972266H1	1602	1872
34	LG:332944.2:2000SEP08	6912541J1	1208	1773
34	LG:332944.2:2000SEP08	4218686H1	1240	1415
34	LG:332944.2:2000SEP08	4137196H1	1267	1558
34	LG:332944.2:2000SEP08	3494764H1	1267	1549
34	LG:332944.2:2000SEP08	7963303H1	1326	1928
34	LG:332944.2:2000SEP08	6036656H1	1330	1818
34	LG:332944.2:2000SEP08	6890168H1	1372	1774
34	LG:332944.2:2000SEP08	5166784H1	1377	1638
34	LG:332944.2:2000SEP08	2823120H1	1388	1682
34	LG:332944.2:2000SEP08	1419848H1	1423	1683
34	LG:332944.2:2000SEP08	g775811	568	658
34	LG:332944.2:2000SEP08	4759703H1	482	750
34	LG:332944.2:2000SEP08	6447507H1	561	673
34	LG:332944.2:2000SEP08	g572841	567	945
34	LG:332944.2:2000SEP08	g669931	567	843
34	LG:332944.2:2000SEP08	g774795	568	774
34	LG:332944.2:2000SEP08	7014128H1	247	630
34	LG:332944.2:2000SEP08	3494458H1	184	246
34	LG:332944.2:2000SEP08	7180711H1	307	866
34	LG:332944.2:2000SEP08	g1524768	407	845
34	LG:332944.2:2000SEP08	3490151H1	454	753
34	LG:332944.2:2000SEP08	3490151F6	454	934
34	LG:332944.2:2000SEP08	7180634H1	1	573
34	LG:332944.2:2000SEP08	7367862H1	3	532
34	LG:332944.2:2000SEP08	3288708H1	43	298
34	LG:332944.2:2000SEP08	7185461H1	43	599
34	LG:332944.2:2000SEP08	2420706H1	2534	2735

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
34	LG:332944.2:2000SEP08	2294013H1	2605	2734
34	LG:332944.2:2000SEP08	2425692H1	2610	2735
34	LG:332944.2:2000SEP08	g3884654	2681	2735
34	LG:332944.2:2000SEP08	2410374H1	2396	2602
34	LG:332944.2:2000SEP08	625035R6	2520	2735
34	LG:332944.2:2000SEP08	g1524707	2405	2742
34	LG:332944.2:2000SEP08	g2062985	2457	2735
34	LG:332944.2:2000SEP08	g824302	2474	2754
34	LG:332944.2:2000SEP08	g565531	2482	2732
34	LG:332944.2:2000SEP08	3398185H1	2514	2729
34	LG:332944.2:2000SEP08	625035H1	2520	2736
34	LG:332944.2:2000SEP08	625035T6	2520	2698
35	LI:347174.5:2000SEP08	2906333H1	1	204
35	LI:347174.5:2000SEP08	5643863H1	1	256
35	LI:347174.5:2000SEP08	7945154J1	20	602
35	LI:347174.5:2000SEP08	534380R7	87	612
35	LI:347174.5:2000SEP08	534380R1	87	680
35	LI:347174.5:2000SEP08	534380H1	87	248
35	LI:347174.5:2000SEP08	534380F1	142	627
35	LI:347174.5:2000SEP08	7211585H2	162	733
35	LI:347174.5:2000SEP08	g7378097	162	579
35	LI:347174.5:2000SEP08	534380T6	216	591
35	LI:347174.5:2000SEP08	7455324H1	217	788
35	LI:347174.5:2000SEP08	5093165H1	243	374
35	LI:347174.5:2000SEP08	7381059H1	234	842
35	LI:347174.5:2000SEP08	1917605H1	288	550
35	LI:347174.5:2000SEP08	g986717	292	654
35	LI:347174.5:2000SEP08	g1151549	341	743
35	LI:347174.5:2000SEP08	7710384J1	568	1133
35	LI:347174.5:2000SEP08	g853547	626	954
35	LI:347174.5:2000SEP08	g866730	626	924
35	LI:347174.5:2000SEP08	7269304H1	641	1118
35	LI:347174.5:2000SEP08	5689745H1	912	1197
35	LI:347174.5:2000SEP08	g2656573	888	1315
35	LI:347174.5:2000SEP08	8025659J1	748	1343
35	LI:347174.5:2000SEP08	8019711J1	786	1340
35	LI:347174.5:2000SEP08	7657968H1	812	1348
35	LI:347174.5:2000SEP08	5090660H1	848	1104
35	LI:347174.5:2000SEP08	3426289F6	873	1310
35	LI:347174.5:2000SEP08	3426289H1	873	1120
35	LI:347174.5:2000SEP08	8019272J1	885	1468
35	LI:347174.5:2000SEP08	7657968J1	701	1366
35	LI:347174.5:2000SEP08	g2896305	644	1060
35	LI:347174.5:2000SEP08	7330740H2	940	1406
35	LI:347174.5:2000SEP08	g853548	945	1294
35	LI:347174.5:2000SEP08	2478255H1	945	1206
35	LI:347174.5:2000SEP08	7242924H1	956	1457
35	LI:347174.5:2000SEP08	5059272H1	979	1239
35	LI:347174.5:2000SEP08	7710384H1	986	1434

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
35	LI:347174.5:2000SEP08	g986626	992	1312
35	LI:347174.5:2000SEP08	g1493306	1013	1315
35	LI:347174.5:2000SEP08	g1126305	1018	1315
35	LI:347174.5:2000SEP08	g865829	1021	1300
35	LI:347174.5:2000SEP08	g3416624	1022	1315
35	LI:347174.5:2000SEP08	1268777H1	1113	1346
35	LI:347174.5:2000SEP08	1268793F6	1113	1366
35	LI:347174.5:2000SEP08	71485473V1	1113	1203
35	LI:347174.5:2000SEP08	1268793H1	1113	1333
35	LI:347174.5:2000SEP08	1268793T6	1117	1325
35	LI:347174.5:2000SEP08	6806077H1	1142	1518
35	LI:347174.5:2000SEP08	7331573H1	1144	1342
36	LI:477070.1:2000SEP08	5906002F8	2	517
36	LI:477070.1:2000SEP08	5906002F6	1	465
36	LI:477070.1:2000SEP08	5906002H1	1	311
36	LI:477070.1:2000SEP08	5906081H1	1	278
36	LI:477070.1:2000SEP08	5906002T9	272	798
37	LI:723144.1:2000SEP08	6274684F8	1	632
37	LI:723144.1:2000SEP08	6274684H1	1	508
37	LI:723144.1:2000SEP08	6274684T8	380	565
38	LI:1007188.1:2000SEP08	6796314H1	1	552
38	LI:1007188.1:2000SEP08	6796314T8	1	534
38	LI:1007188.1:2000SEP08	6796314F8	1	495
39	LI:1024412.1:2000SEP08	6797843F8	1	436
39	LI:1024412.1:2000SEP08	6797843H1	1	436
39	LI:1024412.1:2000SEP08	6797843T8	1	334
40	LI:284797.3:2000SEP08	6930643H1	1	318
40	LI:284797.3:2000SEP08	7388278H1	174	617
40	LI:284797.3:2000SEP08	6992433H1	227	467
40	LI:284797.3:2000SEP08	382351H1	243	498
40	LI:284797.3:2000SEP08	382351R6	243	442
40	LI:284797.3:2000SEP08	6294383H1	317	546
40	LI:284797.3:2000SEP08	6295384H1	317	584
40	LI:284797.3:2000SEP08	6291724H1	318	537
40	LI:284797.3:2000SEP08	6292792H1	317	524
40	LI:284797.3:2000SEP08	7055858H1	449	729
40	LI:284797.3:2000SEP08	7109844H1	513	1015
40	LI:284797.3:2000SEP08	5901582H1	530	811
40	LI:284797.3:2000SEP08	5901582F6	530	965
40	LI:284797.3:2000SEP08	6989086H1	780	1292
40	LI:284797.3:2000SEP08	3331102H1	1087	1341
40	LI:284797.3:2000SEP08	5901582T6	1162	1681
40	LI:284797.3:2000SEP08	382351T6	1173	1645
40	LI:284797.3:2000SEP08	g1421943	1287	1666
40	LI:284797.3:2000SEP08	5901140T6	1299	1646
40	LI:284797.3:2000SEP08	7409070H1	1441	2067
40	LI:284797.3:2000SEP08	6990104H1	1575	2036
41	LI:1092901.1:2000SEP08	g2884969	492	892
41	LI:1092901.1:2000SEP08	6702215H1	1	640



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
42	LI:228930.1:2000SEP08	g5676063	281	752
42	LI:228930.1:2000SEP08	1576986F6	300	691
42	LI:228930.1:2000SEP08	1576986H1	300	379
42	LI:228930.1:2000SEP08	1576986T6	305	713
42	LI:228930.1:2000SEP08	g3843718	394	754
42	LI:228930.1:2000SEP08	g3750491	438	752
42	LI:228930.1:2000SEP08	3604635H1	439	614
42	LI:228930.1:2000SEP08	4753733F8	152	643
42	LI:228930.1:2000SEP08	7065882H1	213	709
42	LI:228930.1:2000SEP08	60127435D1	1	373
42	LI:228930.1:2000SEP08	60127438D1	1	395
42	LI:228930.1:2000SEP08	6991066H1	31	220
42	LI:228930.1:2000SEP08	4753733H1	151	309
42	LI:228930.1:2000SEP08	60127438D4	1	342
43	LI:722913.1:2000SEP08	6273688F8	1	691
43	LI:722913.1:2000SEP08	6273163H1	1	514
43	LI:722913.1:2000SEP08	6273163T8	391	928
44	LG:457478.1:2000SEP08	6850916H1	1	561
44	LG:457478.1:2000SEP08	g4509558	342	569
45	LG:358719.1:2000SEP08	3589295H1	249	542
45	LG:358719.1:2000SEP08	3590122H1	1	310
45	LG:358719.1:2000SEP08	3315281H1	83	339
45	LG:358719.1:2000SEP08	3585294F6	103	524
45	LG:358719.1:2000SEP08	3585294H1	103	320
45	LG:358719.1:2000SEP08	3461069H1	246	484
45	LG:358719.1:2000SEP08	3462962H1	248	364
45	LG:358719.1:2000SEP08	3460259H1	248	479
45	LG:358719.1:2000SEP08	5674847H1	249	504
45	LG:358719.1:2000SEP08	3591414H1	249	560
45	LG:358719.1:2000SEP08	3585880H1	249	347
45	LG:358719.1:2000SEP08	3592746H1	249	547
45	LG:358719.1:2000SEP08	5674646F6	249	546
45	LG:358719.1:2000SEP08	3588813H1	249	565
45	LG:358719.1:2000SEP08	5675060H1	249	504
45	LG:358719.1:2000SEP08	3593262H1	249	571
45	LG:358719.1:2000SEP08	5674646H1	249	503
45	LG:358719.1:2000SEP08	3461227H1	249	488
45	LG:358719.1:2000SEP08	3591576H1	250	552
45	LG:358719.1:2000SEP08	3315732F6	250	736
45	LG:358719.1:2000SEP08	3580084H1	251	403
45	LG:358719.1:2000SEP08	3315732H1	251	488
45	LG:358719.1:2000SEP08	3592541H1	252	554
45	LG:358719.1:2000SEP08	3458495H1	252	497
45	LG:358719.1:2000SEP08	3568588H1	277	469
45	LG:358719.1:2000SEP08	7274542H1	277	757
45	LG:358719.1:2000SEP08	3592638H1	277	544
45	LG:358719.1:2000SEP08	7274539H1	277	728
45	LG:358719.1:2000SEP08	3459843H1	277	498
45	LG:358719.1:2000SEP08	3593421H1	277	569

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
45	LG:358719.1:2000SEP08	5674306H1	362	597
45	LG:358719.1:2000SEP08	5674406H1	363	598
45	LG:358719.1:2000SEP08	5674646T6	685	1185
45	LG:358719.1:2000SEP08	3315732T6	686	1247
45	LG:358719.1:2000SEP08	3585294T6	909	1209
46	LG:105160.5:2000SEP08	4069650T6	488	1017
46	LG:105160.5:2000SEP08	4073607T6	488	1017
46	LG:105160.5:2000SEP08	5615211R6	621	786
46	LG:105160.5:2000SEP08	5614862R8	282	785
46	LG:105160.5:2000SEP08	g316172	168	522
46	LG:105160.5:2000SEP08	7639379J1	1	472
47	LG:400705.1:2000SEP08	g4618234	25	161
47	LG:400705.1:2000SEP08	g3658871	13	448
47	LG:400705.1:2000SEP08	6758610J1	1	219
47	LG:400705.1:2000SEP08	2868968H1	846	996
47	LG:400705.1:2000SEP08	2045886H1	893	975
47	LG:400705.1:2000SEP08	3841273F6	943	996
47	LG:400705.1:2000SEP08	2685577F6	575	996
47	LG:400705.1:2000SEP08	g5857600	634	996
47	LG:400705.1:2000SEP08	2685577H1	575	830
47	LG:400705.1:2000SEP08	2872218H1	603	880
47	LG:400705.1:2000SEP08	1291082H1	775	1004
47	LG:400705.1:2000SEP08	g4435272	626	971
47	LG:400705.1:2000SEP08	2685577T6	568	956
47	LG:400705.1:2000SEP08	g6398653	572	997
47	LG:400705.1:2000SEP08	3438037H1	322	539
47	LG:400705.1:2000SEP08	6035629H1	455	514
47	LG:400705.1:2000SEP08	7277253H1	468	979
48	LG:221977.1:2000SEP08	2553218F6	1	403
48	LG:221977.1:2000SEP08	2433120H1	1	268
48	LG:221977.1:2000SEP08	2828575H1	1	261
48	LG:221977.1:2000SEP08	4738332F6	1	447
48	LG:221977.1:2000SEP08	2553218H1	1	238
48	LG:221977.1:2000SEP08	3165349H1	1	292
48	LG:221977.1:2000SEP08	5085743H1	1	250
48	LG:221977.1:2000SEP08	2434813H1	1	228
48	LG:221977.1:2000SEP08	656611H1	1	252
48	LG:221977.1:2000SEP08	637721F1	1326	1631
48	LG:221977.1:2000SEP08	504384H1	1325	1528
48	LG:221977.1:2000SEP08	g3777925	1327	1629
48	LG:221977.1:2000SEP08	2437205H1	1329	1570
48	LG:221977.1:2000SEP08	g2789160	1327	1626
48	LG:221977.1:2000SEP08	g2767818	1332	1626
48	LG:221977.1:2000SEP08	1878166T6	1333	1588
48	LG:221977.1:2000SEP08	343180H1	1334	1547
48	LG:221977.1:2000SEP08	g3040984	1334	1627
48	LG:221977.1:2000SEP08	g5445738	1373	1622
48	LG:221977.1:2000SEP08	g3042705	1373	1622
48	LG:221977.1:2000SEP08	4367132H1	1381	1650

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
48	LG:221977.1:2000SEP08	5079084H1	1410	1620
48	LG:221977.1:2000SEP08	5337496H1	1420	1622
48	LG:221977.1:2000SEP08	4673726H1	1443	1626
48	LG:221977.1:2000SEP08	6543758H1	1450	1642
48	LG:221977.1:2000SEP08	g4189663	1508	1623
48	LG:221977.1:2000SEP08	g4001918	1508	1621
48	LG:221977.1:2000SEP08	g4071812	1520	1622
48	LG:221977.1:2000SEP08	g3321455	1522	1622
48	LG:221977.1:2000SEP08	g4004309	1522	1622
48	LG:221977.1:2000SEP08	g4002213	1522	1621
48	LG:221977.1:2000SEP08	g4086907	1522	1616
48	LG:221977.1:2000SEP08	g4086502	1531	1626
48	LG:221977.1:2000SEP08	g4071826	1530	1621
48	LG:221977.1:2000SEP08	g4071827	1531	1626
48	LG:221977.1:2000SEP08	g3806277	1531	1621
48	LG:221977.1:2000SEP08	g4077659	1531	1622
48	LG:221977.1:2000SEP08	g3890415	1531	1622
48	LG:221977.1:2000SEP08	g4071856	1531	1626
48	LG:221977.1:2000SEP08	g4018031	1531	1628
48	LG:221977.1:2000SEP08	g4086362	1531	1622
48	LG:221977.1:2000SEP08	g3807025	1531	1619
48	LG:221977.1:2000SEP08	g3847087	1531	1623
48	LG:221977.1:2000SEP08	g3806490	1531	1616
48	LG:221977.1:2000SEP08	g4001960	1531	1622
48	LG:221977.1:2000SEP08	g4077344	1531	1623
48	LG:221977.1:2000SEP08	2467803H1	1546	1624
48	LG:221977.1:2000SEP08	2767324F6	1546	1627
48	LG:221977.1:2000SEP08	2767324H1	1546	1631
48	LG:221977.1:2000SEP08	2655076H1	1551	1629
48	LG:221977.1:2000SEP08	1877044H1	1038	1293
48	LG:221977.1:2000SEP08	7741329H1	1046	1183
48	LG:221977.1:2000SEP08	7741329J1	1048	1183
48	LG:221977.1:2000SEP08	5714566H1	1059	1336
48	LG:221977.1:2000SEP08	5088294H1	1104	1338
48	LG:221977.1:2000SEP08	g4372183	1165	1632
48	LG:221977.1:2000SEP08	g5233177	1175	1632
48	LG:221977.1:2000SEP08	473745R1	979	1500
48	LG:221977.1:2000SEP08	2649470H1	985	1172
48	LG:221977.1:2000SEP08	4349739H1	985	1172
48	LG:221977.1:2000SEP08	5602714H1	982	1179
48	LG:221977.1:2000SEP08	2022038H1	984	1172
48	LG:221977.1:2000SEP08	1873520H1	1002	1172
48	LG:221977.1:2000SEP08	6569366H1	1008	1435
48	LG:221977.1:2000SEP08	5080171H1	1016	1249
48	LG:221977.1:2000SEP08	1878166H1	1021	1264
48	LG:221977.1:2000SEP08	6516938H1	1034	1500
48	LG:221977.1:2000SEP08	1881360H1	978	1172
48	LG:221977.1:2000SEP08	1898227H2	939	1187
48	LG:221977.1:2000SEP08	3509093H1	940	1172

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
48	LG:221977.1:2000SEP08	2331493H1	972	1172
48	LG:221977.1:2000SEP08	2331493R6	972	1467
48	LG:221977.1:2000SEP08	g2079336	834	1060
48	LG:221977.1:2000SEP08	6566590H1	850	1367
48	LG:221977.1:2000SEP08	2556102H1	978	1172
48	LG:221977.1:2000SEP08	4631933H1	889	1163
48	LG:221977.1:2000SEP08	5606960H1	904	1155
48	LG:221977.1:2000SEP08	5082533H1	1	242
48	LG:221977.1:2000SEP08	658193H1	1	223
48	LG:221977.1:2000SEP08	1878854H1	1	270
48	LG:221977.1:2000SEP08	4738332H2	1	224
48	LG:221977.1:2000SEP08	4957834H1	3	245
48	LG:221977.1:2000SEP08	655627H1	1	277
48	LG:221977.1:2000SEP08	4045716F8	11	571
48	LG:221977.1:2000SEP08	4045716H1	12	309
48	LG:221977.1:2000SEP08	5085743F8	16	267
48	LG:221977.1:2000SEP08	4738332R6	69	457
48	LG:221977.1:2000SEP08	7336580H1	283	660
48	LG:221977.1:2000SEP08	7336680H1	290	840
48	LG:221977.1:2000SEP08	2434183H1	601	801
48	LG:221977.1:2000SEP08	638052H1	740	1002
48	LG:221977.1:2000SEP08	637721H1	740	1001
48	LG:221977.1:2000SEP08	637721R1	740	1294
48	LG:221977.1:2000SEP08	4347308H1	770	1023
48	LG:221977.1:2000SEP08	5465601H1	807	1059
48	LG:221977.1:2000SEP08	3939403H1	808	1073
48	LG:221977.1:2000SEP08	3214389H1	823	1039
48	LG:221977.1:2000SEP08	4570745H1	824	1085
48	LG:221977.1:2000SEP08	5466777H1	832	1057
48	LG:221977.1:2000SEP08	2830536H1	832	1091
48	LG:221977.1:2000SEP08	g2208480	1323	1622
48	LG:221977.1:2000SEP08	2753928H1	1291	1556
48	LG:221977.1:2000SEP08	4630914H1	1288	1530
48	LG:221977.1:2000SEP08	656173H1	1291	1556
48	LG:221977.1:2000SEP08	6156346H1	1291	1581
48	LG:221977.1:2000SEP08	3180154H1	1293	1596
48	LG:221977.1:2000SEP08	1343224H1	1294	1516
48	LG:221977.1:2000SEP08	5080293H1	1293	1467
48	LG:221977.1:2000SEP08	5713922H1	1293	1584
48	LG:221977.1:2000SEP08	3105114H1	1293	1589
48	LG:221977.1:2000SEP08	3868973H1	1293	1570
48	LG:221977.1:2000SEP08	4664924H1	1293	1560
48	LG:221977.1:2000SEP08	3190924H1	1294	1618
48	LG:221977.1:2000SEP08	210395H1	1285	1470
48	LG:221977.1:2000SEP08	5713168H1	1288	1590
48	LG:221977.1:2000SEP08	737775H1	1288	1513
48	LG:221977.1:2000SEP08	2654368H1	1287	1506
48	LG:221977.1:2000SEP08	4803231H1	1290	1555
48	LG:221977.1:2000SEP08	4741765H1	1291	1575

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
48	LG:221977.1:2000SEP08	5688001H1	1288	1564
48	LG:221977.1:2000SEP08	6545268H1	1285	1622
48	LG:221977.1:2000SEP08	4189426H1	1284	1508
48	LG:221977.1:2000SEP08	1839823H1	1279	1549
48	LG:221977.1:2000SEP08	207209H1	1285	1438
48	LG:221977.1:2000SEP08	5712106H1	1280	1581
48	LG:221977.1:2000SEP08	g2873724	1279	1622
48	LG:221977.1:2000SEP08	2767512H1	1284	1541
48	LG:221977.1:2000SEP08	2331493T6	1197	1585
48	LG:221977.1:2000SEP08	2553241H1	1197	1392
48	LG:221977.1:2000SEP08	3798071H1	1197	1421
48	LG:221977.1:2000SEP08	g5369458	1203	1622
48	LG:221977.1:2000SEP08	g6030746	1205	1623
48	LG:221977.1:2000SEP08	1881461T6	1205	1584
48	LG:221977.1:2000SEP08	g6030939	1209	1623
48	LG:221977.1:2000SEP08	1881461F6	1212	1625
48	LG:221977.1:2000SEP08	1881461H1	1212	1489
48	LG:221977.1:2000SEP08	5463734H1	1218	1462
48	LG:221977.1:2000SEP08	3744608H1	1230	1529
48	LG:221977.1:2000SEP08	g2874931	1254	1626
48	LG:221977.1:2000SEP08	3160484H1	1276	1544
48	LG:221977.1:2000SEP08	3159648H1	1276	1531
48	LG:221977.1:2000SEP08	5336872H1	1276	1517
48	LG:221977.1:2000SEP08	2751112H1	1276	1371
48	LG:221977.1:2000SEP08	2654442H1	1276	1555
48	LG:221977.1:2000SEP08	705530R6	1278	1622
48	LG:221977.1:2000SEP08	763848H1	1278	1453
48	LG:221977.1:2000SEP08	3178935H1	1277	1542
48	LG:221977.1:2000SEP08	6546959H1	1278	1642
48	LG:221977.1:2000SEP08	3181192H1	1277	1588
48	LG:221977.1:2000SEP08	705530T6	1278	1583
48	LG:221977.1:2000SEP08	705530H1	1278	1496
48	LG:221977.1:2000SEP08	5078477H1	1278	1510
48	LG:221977.1:2000SEP08	4846582H1	1278	1533
48	LG:221977.1:2000SEP08	6322778H1	1279	1528
49	LG:898771.1:2000SEP08	g2162199	1987	2190
49	LG:898771.1:2000SEP08	7281807H1	2137	2364
49	LG:898771.1:2000SEP08	6600395H1	1911	2323
49	LG:898771.1:2000SEP08	5626829H1	1987	2258
49	LG:898771.1:2000SEP08	g4081510	1	339
49	LG:898771.1:2000SEP08	6862186H1	1	285
49	LG:898771.1:2000SEP08	6766056J1	1	602
49	LG:898771.1:2000SEP08	g5543132	20	418
49	LG:898771.1:2000SEP08	g3246528	20	402
49	LG:898771.1:2000SEP08	g5636120	20	461
49	LG:898771.1:2000SEP08	g4196555	20	300
49	LG:898771.1:2000SEP08	g5636142	20	497
49	LG:898771.1:2000SEP08	g2568732	54	503
49	LG:898771.1:2000SEP08	3747228H1	68	366

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
49	LG:898771.1:2000SEP08	2240520F6	120	446
49	LG:898771.1:2000SEP08	2240520H1	120	221
49	LG:898771.1:2000SEP08	6862047H1	178	457
49	LG:898771.1:2000SEP08	g2209650	294	787
49	LG:898771.1:2000SEP08	6216076H1	320	801
49	LG:898771.1:2000SEP08	7712493J1	396	857
49	LG:898771.1:2000SEP08	7712493H1	396	857
49	LG:898771.1:2000SEP08	7039186H1	419	987
49	LG:898771.1:2000SEP08	6978931R8	464	1158
49	LG:898771.1:2000SEP08	g2209760	571	1048
49	LG:898771.1:2000SEP08	6895681R8	914	1471
49	LG:898771.1:2000SEP08	6895723R8	935	1265
49	LG:898771.1:2000SEP08	6978931F8	1103	1668
49	LG:898771.1:2000SEP08	6978931H1	1317	1666
49	LG:898771.1:2000SEP08	7204172H1	1325	1860
49	LG:898771.1:2000SEP08	7714080H1	1623	2053
49	LG:898771.1:2000SEP08	6819816H1	1692	1890
49	LG:898771.1:2000SEP08	6836688H1	1707	2284
49	LG:898771.1:2000SEP08	6349781H2	1758	1940
49	LG:898771.1:2000SEP08	7280863H1	1801	2204
49	LG:898771.1:2000SEP08	6210267H1	2195	2297
49	LG:898771.1:2000SEP08	7238641H1	2189	2303
50	LI:457478.1:2000SEP08	g1618619	1	235
50	LI:457478.1:2000SEP08	6850916H1	1	563
50	LI:457478.1:2000SEP08	2051259H1	81	365
50	LI:457478.1:2000SEP08	3036424H1	161	446
50	LI:457478.1:2000SEP08	g2356855	180	538
50	LI:457478.1:2000SEP08	g4509558	344	571
51	LI:125140.1:2000SEP08	g6398056	1	248
51	LI:125140.1:2000SEP08	7644338H1	1	412
51	LI:125140.1:2000SEP08	4434081F6	329	813
51	LI:125140.1:2000SEP08	4434081H1	330	600
51	LI:125140.1:2000SEP08	3371506H1	479	638
51	LI:125140.1:2000SEP08	6832669H1	586	1146
51	LI:125140.1:2000SEP08	7699223J1	636	1233
51	LI:125140.1:2000SEP08	1998681H1	710	964
51	LI:125140.1:2000SEP08	7644338J1	846	1410
52	LI:021095.2:2000SEP08	3232873H1	1	271
53	LI:888730.1:2000SEP08	6798649H1	3	303
53	LI:888730.1:2000SEP08	6790981F8	1	502
53	LI:888730.1:2000SEP08	6795715F8	3	576
53	LI:888730.1:2000SEP08	6795715H1	3	515
53	LI:888730.1:2000SEP08	6798649F8	3	493
53	LI:888730.1:2000SEP08	8081234H2	410	934
53	LI:888730.1:2000SEP08	6798649T8	604	792
53	LI:888730.1:2000SEP08	8081303H2	337	874
53	LI:888730.1:2000SEP08	g1614937	615	921
53	LI:888730.1:2000SEP08	6793426T8	667	798
53	LI:888730.1:2000SEP08	g1616171	724	919

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
53	LI:888730.1:2000SEP08	g1616172	741	920
53	LI:888730.1:2000SEP08	6793426H1	14	459
53	LI:888730.1:2000SEP08	6790981T8	173	811
53	LI:888730.1:2000SEP08	6790981H1	1	491
54	LI:358719.1:2000SEP08	70254766V1	302	872
54	LI:358719.1:2000SEP08	70257814V1	770	1040
54	LI:358719.1:2000SEP08	70255598V1	803	1324
54	LI:358719.1:2000SEP08	71501194V1	492	655
54	LI:358719.1:2000SEP08	71649710V1	565	850
54	LI:358719.1:2000SEP08	71659859V1	570	969
54	LI:358719.1:2000SEP08	70251077V1	507	705
54	LI:358719.1:2000SEP08	70246852V1	514	671
54	LI:358719.1:2000SEP08	70255542V1	524	1027
54	LI:358719.1:2000SEP08	70249977V1	528	722
54	LI:358719.1:2000SEP08	71651191V1	1110	1308
54	LI:358719.1:2000SEP08	71644160V1	768	1311
54	LI:358719.1:2000SEP08	70257007V2	760	1125
54	LI:358719.1:2000SEP08	70254569V1	876	1297
54	LI:358719.1:2000SEP08	71497489V1	889	1078
54	LI:358719.1:2000SEP08	71497304V1	889	1106
54	LI:358719.1:2000SEP08	70249846V1	908	1077
54	LI:358719.1:2000SEP08	3585294T6	948	1253
54	LI:358719.1:2000SEP08	70251447V1	1010	1097
54	LI:358719.1:2000SEP08	70251268V1	1028	1175
54	LI:358719.1:2000SEP08	70247332V1	1086	1455
54	LI:358719.1:2000SEP08	70250954V1	1104	1289
54	LI:358719.1:2000SEP08	70254862V1	282	832
54	LI:358719.1:2000SEP08	3568588H1	283	487
54	LI:358719.1:2000SEP08	7274542H1	283	793
54	LI:358719.1:2000SEP08	70248496V1	719	1128
54	LI:358719.1:2000SEP08	70248691V1	719	1128
54	LI:358719.1:2000SEP08	70257007V1	731	1125
54	LI:358719.1:2000SEP08	70254747V1	780	1304
54	LI:358719.1:2000SEP08	3591576H1	254	572
54	LI:358719.1:2000SEP08	70254628V1	254	744
54	LI:358719.1:2000SEP08	3592541H1	256	574
54	LI:358719.1:2000SEP08	71686858V1	268	464
54	LI:358719.1:2000SEP08	5674406H1	376	623
54	LI:358719.1:2000SEP08	71639758V1	417	1008
54	LI:358719.1:2000SEP08	70256300V1	813	1328
54	LI:358719.1:2000SEP08	70257439V1	811	1015
54	LI:358719.1:2000SEP08	70254995V1	814	1297
54	LI:358719.1:2000SEP08	3588813H1	253	585
54	LI:358719.1:2000SEP08	3593262H1	253	593
54	LI:358719.1:2000SEP08	3315732H1	255	507
54	LI:358719.1:2000SEP08	3315281H1	83	350
54	LI:358719.1:2000SEP08	71496461V1	103	645
54	LI:358719.1:2000SEP08	71499658V1	103	700
54	LI:358719.1:2000SEP08	71500936V1	103	829

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
54	LI:358719.1:2000SEP08	3585294F6	103	544
54	LI:358719.1:2000SEP08	70247644V1	584	837
54	LI:358719.1:2000SEP08	71645502V1	620	1030
54	LI:358719.1:2000SEP08	71642987V1	615	1337
54	LI:358719.1:2000SEP08	70258411V1	664	880
54	LI:358719.1:2000SEP08	71666717V1	680	977
54	LI:358719.1:2000SEP08	5674646T6	718	1228
54	LI:358719.1:2000SEP08	3315732T6	719	1291
54	LI:358719.1:2000SEP08	71642544V1	722	1259
54	LI:358719.1:2000SEP08	70249638V1	358	919
54	LI:358719.1:2000SEP08	70256178V1	254	701
54	LI:358719.1:2000SEP08	70249314V1	444	531
54	LI:358719.1:2000SEP08	71684549V1	472	718
54	LI:358719.1:2000SEP08	71641793V1	253	882
54	LI:358719.1:2000SEP08	71644835V1	250	813
54	LI:358719.1:2000SEP08	3461069H1	250	503
54	LI:358719.1:2000SEP08	71686758V1	249	463
54	LI:358719.1:2000SEP08	70249514V1	1103	1265
54	LI:358719.1:2000SEP08	71649473V1	1120	1205
54	LI:358719.1:2000SEP08	70256285V1	254	733
54	LI:358719.1:2000SEP08	70257039V2	254	628
54	LI:358719.1:2000SEP08	3593421H1	283	591
54	LI:358719.1:2000SEP08	3592638H1	283	564
54	LI:358719.1:2000SEP08	3459843H1	283	516
54	LI:358719.1:2000SEP08	70255330V1	292	787
54	LI:358719.1:2000SEP08	71501490V1	301	800
54	LI:358719.1:2000SEP08	71526655V1	425	607
54	LI:358719.1:2000SEP08	3590122H1	1	317
54	LI:358719.1:2000SEP08	70506883V1	825	1297
54	LI:358719.1:2000SEP08	3460259H1	252	497
54	LI:358719.1:2000SEP08	71640275V1	253	800
54	LI:358719.1:2000SEP08	3591414H1	253	580
54	LI:358719.1:2000SEP08	5674646F6	253	566
54	LI:358719.1:2000SEP08	3462962H1	252	377
54	LI:358719.1:2000SEP08	3461227H1	253	507
54	LI:358719.1:2000SEP08	3585880H1	253	358
54	LI:358719.1:2000SEP08	70256157V1	254	830
54	LI:358719.1:2000SEP08	70255182V1	254	808
54	LI:358719.1:2000SEP08	3315732F6	254	772
54	LI:358719.1:2000SEP08	70255417V1	846	1324
54	LI:358719.1:2000SEP08	70255818V1	860	1315
54	LI:358719.1:2000SEP08	70254446V1	869	1297
54	LI:358719.1:2000SEP08	70254414V1	870	1274
54	LI:358719.1:2000SEP08	70251204V1	254	349
54	LI:358719.1:2000SEP08	7274539H1	283	763
54	LI:358719.1:2000SEP08	3592746H1	253	567
54	LI:358719.1:2000SEP08	3589295H1	253	562
54	LI:358719.1:2000SEP08	70255284V1	254	583
54	LI:358719.1:2000SEP08	5674847H1	253	522



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
54	U:358719.1:2000SEP08	5674646H1	253	521
54	U:358719.1:2000SEP08	5675060H1	253	522
54	U:358719.1:2000SEP08	70258102V1	432	661
54	U:358719.1:2000SEP08	3580084H1	255	417
54	U:358719.1:2000SEP08	70255950V1	254	801
54	U:358719.1:2000SEP08	3458495H1	256	515
54	U:358719.1:2000SEP08	71498511V1	344	785
54	U:358719.1:2000SEP08	70255091V1	270	736
54	U:358719.1:2000SEP08	3585294H1	103	330
54	U:358719.1:2000SEP08	71497471V1	149	588
54	U:358719.1:2000SEP08	71471279V1	215	524
54	U:358719.1:2000SEP08	70254882V1	370	942
54	U:358719.1:2000SEP08	5674306H1	375	622
54	U:358719.1:2000SEP08	70255110V1	254	817
55	U:351342.3:2000SEP08	6819816H1	819	1017
55	U:351342.3:2000SEP08	6836688H1	425	1002
55	U:351342.3:2000SEP08	7280863H1	505	908
55	U:351342.3:2000SEP08	6600395H1	386	798
55	U:351342.3:2000SEP08	5626829H1	451	722
55	U:351342.3:2000SEP08	g2162199	519	722
55	U:351342.3:2000SEP08	7281807H1	352	572
55	U:351342.3:2000SEP08	7238641H1	405	520
55	U:351342.3:2000SEP08	7238641F8	1	517
55	U:351342.3:2000SEP08	6210267H1	412	514
55	U:351342.3:2000SEP08	g4081510	1137	1484
55	U:351342.3:2000SEP08	6862186H1	1191	1477
55	U:351342.3:2000SEP08	8019650J1	874	1475
55	U:351342.3:2000SEP08	8019607J1	949	1475
55	U:351342.3:2000SEP08	6766056J1	874	1475
55	U:351342.3:2000SEP08	8053889J1	1108	1475
55	U:351342.3:2000SEP08	g3246528	1074	1456
55	U:351342.3:2000SEP08	g5636120	1015	1456
55	U:351342.3:2000SEP08	g5636142	979	1456
55	U:351342.3:2000SEP08	g4196555	1176	1456
55	U:351342.3:2000SEP08	g2568732	973	1422
55	U:351342.3:2000SEP08	7204172H1	849	1384
55	U:351342.3:2000SEP08	2240520F6	1030	1356
55	U:351342.3:2000SEP08	7714080H1	656	1086
56	U:256099.2:2000SEP08	71620166V1	342	1039
56	U:256099.2:2000SEP08	4399381H1	479	775
56	U:256099.2:2000SEP08	71615821V1	342	1108
56	U:256099.2:2000SEP08	71617264V1	721	1501
56	U:256099.2:2000SEP08	71615884V1	722	1470
56	U:256099.2:2000SEP08	71596772V1	677	922
56	U:256099.2:2000SEP08	71132277V1	682	1355
56	U:256099.2:2000SEP08	70191964V1	341	876
56	U:256099.2:2000SEP08	4623607H1	342	492
56	U:256099.2:2000SEP08	71617431V1	342	1050
56	U:256099.2:2000SEP08	71615865V1	342	1047

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
56	LI:256099.2:2000SEP08	70191812V1	620	1231
56	LI:256099.2:2000SEP08	71621090V1	627	765
56	LI:256099.2:2000SEP08	496221H1	176	404
56	LI:256099.2:2000SEP08	71191285V1	1222	1871
56	LI:256099.2:2000SEP08	71616554V1	466	1208
56	LI:256099.2:2000SEP08	71132568V1	1227	1702
56	LI:256099.2:2000SEP08	71129593V1	1232	1446
56	LI:256099.2:2000SEP08	2721125H1	1237	1497
56	LI:256099.2:2000SEP08	71619624V1	1249	1895
56	LI:256099.2:2000SEP08	70191992V1	389	925
56	LI:256099.2:2000SEP08	3774992H1	389	721
56	LI:256099.2:2000SEP08	6934512H1	1368	1878
56	LI:256099.2:2000SEP08	71132775V1	1371	1876
56	LI:256099.2:2000SEP08	220732R1	811	1438
56	LI:256099.2:2000SEP08	71131234V1	817	1443
56	LI:256099.2:2000SEP08	71304572V1	1267	1599
56	LI:256099.2:2000SEP08	71131838V1	48	351
56	LI:256099.2:2000SEP08	2174518H1	51	315
56	LI:256099.2:2000SEP08	3572842H1	78	369
56	LI:256099.2:2000SEP08	g1561727	80	1907
56	LI:256099.2:2000SEP08	876929H1	235	348
56	LI:256099.2:2000SEP08	71616214V1	1216	1880
56	LI:256099.2:2000SEP08	71190402V1	994	1648
56	LI:256099.2:2000SEP08	71191544V1	996	1630
56	LI:256099.2:2000SEP08	70192173V1	1000	1583
56	LI:256099.2:2000SEP08	71131089V1	152	499
56	LI:256099.2:2000SEP08	3947429H1	151	449
56	LI:256099.2:2000SEP08	70191815V1	807	1341
56	LI:256099.2:2000SEP08	220732H1	812	1039
56	LI:256099.2:2000SEP08	71617637V1	1204	1895
56	LI:256099.2:2000SEP08	70193242V1	1224	1508
56	LI:256099.2:2000SEP08	71191065V1	151	692
56	LI:256099.2:2000SEP08	71188528V1	151	829
56	LI:256099.2:2000SEP08	71620277V1	768	1449
56	LI:256099.2:2000SEP08	71620252V1	768	1448
56	LI:256099.2:2000SEP08	70192498V1	342	797
56	LI:256099.2:2000SEP08	3739474H1	501	823
56	LI:256099.2:2000SEP08	3383537H1	576	841
56	LI:256099.2:2000SEP08	71657051V1	589	1003
56	LI:256099.2:2000SEP08	70192196V1	602	1136
56	LI:256099.2:2000SEP08	71620637V1	342	1049
56	LI:256099.2:2000SEP08	71618785V1	342	980
56	LI:256099.2:2000SEP08	71129139V1	1183	1325
56	LI:256099.2:2000SEP08	71616871V1	1204	1897
56	LI:256099.2:2000SEP08	g1384884	1412	1892
56	LI:256099.2:2000SEP08	71597459V1	955	1716
56	LI:256099.2:2000SEP08	71622214V1	982	1359
56	LI:256099.2:2000SEP08	5168528H1	970	1301
56	LI:256099.2:2000SEP08	71190616V1	151	721

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
56	LI:256099.2:2000SEP08	g1616070	183	335
56	LI:256099.2:2000SEP08	7710212H1	198	761
56	LI:256099.2:2000SEP08	6947589H1	217	785
56	LI:256099.2:2000SEP08	832562R1	106	611
56	LI:256099.2:2000SEP08	71133469V1	1024	1470
56	LI:256099.2:2000SEP08	6538702H1	1047	1447
56	LI:256099.2:2000SEP08	2757434H1	1081	1377
56	LI:256099.2:2000SEP08	71619911V1	1012	1710
56	LI:256099.2:2000SEP08	71191068V1	1017	1374
56	LI:256099.2:2000SEP08	70191312V1	342	861
56	LI:256099.2:2000SEP08	1461274H1	492	767
56	LI:256099.2:2000SEP08	71129244V1	520	1195
56	LI:256099.2:2000SEP08	70191690V1	342	782
56	LI:256099.2:2000SEP08	5529378H1	1313	1558
56	LI:256099.2:2000SEP08	g1388447	1304	1623
56	LI:256099.2:2000SEP08	71132084V1	1280	1850
56	LI:256099.2:2000SEP08	220732F1	1279	1891
56	LI:256099.2:2000SEP08	71130570V1	1172	1308
56	LI:256099.2:2000SEP08	71190752V1	1165	1476
56	LI:256099.2:2000SEP08	71618692V1	342	968
56	LI:256099.2:2000SEP08	70194001V1	1374	1714
56	LI:256099.2:2000SEP08	71129545V1	1383	1891
56	LI:256099.2:2000SEP08	70191322V1	1410	1843
56	LI:256099.2:2000SEP08	71618890V1	1426	1915
56	LI:256099.2:2000SEP08	70192505V1	1426	1891
56	LI:256099.2:2000SEP08	71188839V1	1428	2011
56	LI:256099.2:2000SEP08	71131985V1	602	1013
56	LI:256099.2:2000SEP08	832562H1	106	213
56	LI:256099.2:2000SEP08	1784941H1	106	349
56	LI:256099.2:2000SEP08	71190661V1	151	751
56	LI:256099.2:2000SEP08	5754860H1	147	651
56	LI:256099.2:2000SEP08	71130477V1	818	1508
56	LI:256099.2:2000SEP08	g1616064	183	514
56	LI:256099.2:2000SEP08	2346852F6	48	551
56	LI:256099.2:2000SEP08	2346852H1	48	286
56	LI:256099.2:2000SEP08	7688858H1	1	469
56	LI:256099.2:2000SEP08	71175881V1	1176	1602
56	LI:256099.2:2000SEP08	g1403337	271	1705
56	LI:256099.2:2000SEP08	2568292H1	1263	1522
56	LI:256099.2:2000SEP08	71188630V1	1165	1477
56	LI:256099.2:2000SEP08	71619837V1	797	1429
56	LI:256099.2:2000SEP08	71621179V1	762	1454
56	LI:256099.2:2000SEP08	832627H1	106	334
56	LI:256099.2:2000SEP08	71191604V1	1267	1866
56	LI:256099.2:2000SEP08	71596620V1	1267	1367
56	LI:256099.2:2000SEP08	71615643V1	737	1448
56	LI:256099.2:2000SEP08	71189450V1	151	729
56	LI:256099.2:2000SEP08	3947429F6	151	638
56	LI:256099.2:2000SEP08	71189442V1	151	751

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
56	LI:256099.2:2000SEP08	71191531V1	151	654
56	LI:256099.2:2000SEP08	71188842V1	151	718
56	LI:256099.2:2000SEP08	71620734V1	1329	1928
56	LI:256099.2:2000SEP08	7688858J1	1326	1818
56	LI:256099.2:2000SEP08	70191412V1	1342	1604
56	LI:256099.2:2000SEP08	7355493H1	423	1112
56	LI:256099.2:2000SEP08	70192354V1	1107	1675
56	LI:256099.2:2000SEP08	71620373V1	829	1585
56	LI:256099.2:2000SEP08	71188152V1	827	1180
56	LI:256099.2:2000SEP08	71617256V1	826	1422
56	LI:256099.2:2000SEP08	71617076V1	848	1437
56	LI:256099.2:2000SEP08	71620139V1	843	1461
56	LI:256099.2:2000SEP08	71617242V1	845	1600
56	LI:256099.2:2000SEP08	71619967V1	846	1493
56	LI:256099.2:2000SEP08	71617491V1	850	1476
56	LI:256099.2:2000SEP08	71155601V1	860	1376
56	LI:256099.2:2000SEP08	3880990H1	880	1195
56	LI:256099.2:2000SEP08	71593959V1	888	1417
56	LI:256099.2:2000SEP08	71618259V1	1154	1650
56	LI:256099.2:2000SEP08	3852851H1	1134	1386
56	LI:256099.2:2000SEP08	7930642H1	1149	1696
56	LI:256099.2:2000SEP08	g1382754	1152	1555
56	LI:256099.2:2000SEP08	71618055V1	1084	1830
56	LI:256099.2:2000SEP08	71131975V1	1306	1521
56	LI:256099.2:2000SEP08	3947429T6	1312	1868
56	LI:256099.2:2000SEP08	g1494168	1330	1496
56	LI:256099.2:2000SEP08	70192164V1	1340	1798
56	LI:256099.2:2000SEP08	4623607T6	1344	1860
56	LI:256099.2:2000SEP08	71621424V1	1362	1996
56	LI:256099.2:2000SEP08	067638H1	1361	1543
56	LI:256099.2:2000SEP08	70191215V1	342	783
56	LI:256099.2:2000SEP08	70191765V1	342	450
56	LI:256099.2:2000SEP08	70191531V1	342	736
56	LI:256099.2:2000SEP08	6874432H1	353	1018
56	LI:256099.2:2000SEP08	5302829H1	89	368
56	LI:256099.2:2000SEP08	5329667H1	102	241
56	LI:256099.2:2000SEP08	71618790V1	1428	1886
56	LI:256099.2:2000SEP08	71130063V1	1437	1894
56	LI:256099.2:2000SEP08	71617089V1	1441	1887
56	LI:256099.2:2000SEP08	71616940V1	1440	2011
56	LI:256099.2:2000SEP08	6434868H1	1453	1897
56	LI:256099.2:2000SEP08	4402160H1	1462	1719
56	LI:256099.2:2000SEP08	g2318383	1463	1896
56	LI:256099.2:2000SEP08	71131110V1	1463	1892
56	LI:256099.2:2000SEP08	g1615965	1469	1893
56	LI:256099.2:2000SEP08	g4899694	1471	1898
56	LI:256099.2:2000SEP08	70192423V1	1472	1886
56	LI:256099.2:2000SEP08	g1382698	1474	1895
56	LI:256099.2:2000SEP08	g4898934	1486	1898

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
56	LI:256099.2:2000SEP08	g4764567	1486	1891
56	LI:256099.2:2000SEP08	70192356V1	1525	1895
56	LI:256099.2:2000SEP08	70192168V1	1615	1977
56	LI:256099.2:2000SEP08	70189416V1	1617	1966
56	LI:256099.2:2000SEP08	g4088016	1623	1982
56	LI:256099.2:2000SEP08	g2526316	1643	1979
56	LI:256099.2:2000SEP08	g3038884	1648	1978
56	LI:256099.2:2000SEP08	g3190559	1651	1971
56	LI:256099.2:2000SEP08	g2357999	1679	1979
56	LI:256099.2:2000SEP08	g2957512	1687	1975
56	LI:256099.2:2000SEP08	g3919970	1734	1972
56	LI:256099.2:2000SEP08	g3038673	1737	1971
56	LI:256099.2:2000SEP08	g5127190	1718	1894
56	LI:256099.2:2000SEP08	g3419278	1732	1898
56	LI:256099.2:2000SEP08	5301701H1	1099	1380
56	LI:256099.2:2000SEP08	3326044H1	1105	1304
56	LI:256099.2:2000SEP08	71129862V1	483	1113
56	LI:256099.2:2000SEP08	1461274R6	492	1045
56	LI:256099.2:2000SEP08	71188265V1	492	865
56	LI:256099.2:2000SEP08	71617405V1	785	1561
56	LI:256099.2:2000SEP08	71620141V1	782	1437
56	LI:256099.2:2000SEP08	5022846H1	787	1100
56	LI:256099.2:2000SEP08	70192060V1	805	1300
56	LI:256099.2:2000SEP08	71621174V1	342	909
56	LI:256099.2:2000SEP08	71621290V1	342	835
56	LI:256099.2:2000SEP08	70191745V1	338	867
56	LI:256099.2:2000SEP08	71617104V1	697	1475
57	LI:2051991.1:2000SEP08	6266511H1	451	1034
57	LI:2051991.1:2000SEP08	g1894422	745	1189
57	LI:2051991.1:2000SEP08	5979987H1	940	1197
57	LI:2051991.1:2000SEP08	71673187V1	1	268
57	LI:2051991.1:2000SEP08	539765R6	2	504
57	LI:2051991.1:2000SEP08	539765T6	2	588
57	LI:2051991.1:2000SEP08	539765H1	2	253
57	LI:2051991.1:2000SEP08	71670718V1	112	639
57	LI:2051991.1:2000SEP08	3534754H1	125	397
57	LI:2051991.1:2000SEP08	7467091H1	369	867
57	LI:2051991.1:2000SEP08	g4260485	1065	1438
57	LI:2051991.1:2000SEP08	6497778H1	1109	1591
58	LG:980769.1:2000SEP08	g5177300	422	868
58	LG:980769.1:2000SEP08	g3041438	688	860
58	LG:980769.1:2000SEP08	g2218499	631	845
58	LG:980769.1:2000SEP08	g2737221	467	849
58	LG:980769.1:2000SEP08	g6703559	525	909
58	LG:980769.1:2000SEP08	g6699872	365	909
58	LG:980769.1:2000SEP08	7386340H1	577	901
58	LG:980769.1:2000SEP08	6763732J1	311	900
58	LG:980769.1:2000SEP08	6986658H1	207	307
58	LG:980769.1:2000SEP08	6777292J1	1	650

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
58	LG:980769.1:2000SEP08	g2574686	551	830
58	LG:980769.1:2000SEP08	6777292H1	214	830
59	LG:332474.3:2000SEP08	3334515H1	24	284
59	LG:332474.3:2000SEP08	4833889F8	60	462
59	LG:332474.3:2000SEP08	4833889F9	60	676
59	LG:332474.3:2000SEP08	4833889H1	60	313
59	LG:332474.3:2000SEP08	7584265H1	62	511
59	LG:332474.3:2000SEP08	6858863H1	130	574
59	LG:332474.3:2000SEP08	3334515F6	24	420
59	LG:332474.3:2000SEP08	2290063H1	1	258
59	LG:332474.3:2000SEP08	6858863F8	132	698
59	LG:332474.3:2000SEP08	g1975674	260	544
59	LG:332474.3:2000SEP08	g1991774	357	673
59	LG:332474.3:2000SEP08	6858863T8	392	1015
59	LG:332474.3:2000SEP08	4833889T8	464	869
59	LG:332474.3:2000SEP08	3334515T6	481	1035
59	LG:332474.3:2000SEP08	5428021F7	521	1069
59	LG:332474.3:2000SEP08	5428021H1	521	776
59	LG:332474.3:2000SEP08	g5848557	739	930
59	LG:332474.3:2000SEP08	g5233487	771	1093
59	LG:332474.3:2000SEP08	4773368H1	945	1088
59	LG:332474.3:2000SEP08	4773368F6	945	1088
60	LG:1087707.1:2000SEP08	4333661H1	428	684
60	LG:1087707.1:2000SEP08	7644435J1	536	618
60	LG:1087707.1:2000SEP08	2205615T6	672	1197
60	LG:1087707.1:2000SEP08	g6397537	782	1231
60	LG:1087707.1:2000SEP08	4150193F6	296	804
60	LG:1087707.1:2000SEP08	4150193H1	296	556
60	LG:1087707.1:2000SEP08	4914013H1	331	585
60	LG:1087707.1:2000SEP08	5044566H1	380	657
60	LG:1087707.1:2000SEP08	5044566F6	380	568
60	LG:1087707.1:2000SEP08	g6661611	1	463
61	LG:415349.1:2000SEP08	g5178218	525	946
61	LG:415349.1:2000SEP08	g4190232	540	1005
61	LG:415349.1:2000SEP08	g6710036	534	989
61	LG:415349.1:2000SEP08	g4852090	530	986
61	LG:415349.1:2000SEP08	g4305494	520	985
61	LG:415349.1:2000SEP08	7617296J1	463	959
61	LG:415349.1:2000SEP08	g6993309	575	1024
61	LG:415349.1:2000SEP08	g5914430	565	1016
61	LG:415349.1:2000SEP08	g3679677	542	1005
61	LG:415349.1:2000SEP08	6991133H1	1	269
61	LG:415349.1:2000SEP08	6991888H1	5	493
61	LG:415349.1:2000SEP08	3101689H1	26	294
61	LG:415349.1:2000SEP08	3101689F6	1	282
61	LG:415349.1:2000SEP08	7355332H1	166	748
62	LG:132420.2:2000SEP08	4435742T9	55	677
62	LG:132420.2:2000SEP08	6560279T8	53	677
62	LG:132420.2:2000SEP08	4435742F8	1	558

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
62	LG:132420.2:2000SEP08	7033828H1	978	1541
62	LG:132420.2:2000SEP08	7033828F8	892	1540
62	LG:132420.2:2000SEP08	7033828R8	533	1135
62	LG:132420.2:2000SEP08	g4986202	523	756
62	LG:132420.2:2000SEP08	5102419T6	239	736
63	LG:394201.1:2000SEP08	g3154681	298	707
63	LG:394201.1:2000SEP08	4764960T9	39	642
63	LG:394201.1:2000SEP08	4762939T9	51	603
63	LG:394201.1:2000SEP08	4764960T8	121	554
63	LG:394201.1:2000SEP08	4764960F8	1	515
63	LG:394201.1:2000SEP08	4762939F9	1	469
63	LG:394201.1:2000SEP08	4764960F9	1	406
63	LG:394201.1:2000SEP08	4764960H1	1	88
64	LG:1060884.1:2000SEP08	3679585H1	1	269
64	LG:1060884.1:2000SEP08	2553777H1	1	69
64	LG:1060884.1:2000SEP08	2598119H1	1	161
64	LG:1060884.1:2000SEP08	2598119F6	1	284
64	LG:1060884.1:2000SEP08	3581046H1	1	287
64	LG:1060884.1:2000SEP08	2553777F6	1	340
64	LG:1060884.1:2000SEP08	3679585F7	4	603
64	LG:1060884.1:2000SEP08	4353994F6	66	600
64	LG:1060884.1:2000SEP08	4353354H1	66	263
64	LG:1060884.1:2000SEP08	4353994H1	66	330
64	LG:1060884.1:2000SEP08	549501H1	104	227
64	LG:1060884.1:2000SEP08	g6661142	139	586
64	LG:1060884.1:2000SEP08	2511803H1	239	469
64	LG:1060884.1:2000SEP08	5989639F8	341	927
64	LG:1060884.1:2000SEP08	5989639H1	344	642
64	LG:1060884.1:2000SEP08	6058657F8	345	892
64	LG:1060884.1:2000SEP08	7674643H2	415	943
64	LG:1060884.1:2000SEP08	5185737H1	495	758
64	LG:1060884.1:2000SEP08	g2526396	533	865
64	LG:1060884.1:2000SEP08	g2013023	533	860
64	LG:1060884.1:2000SEP08	4217509H1	553	697
64	LG:1060884.1:2000SEP08	4217509F6	553	1019
64	LG:1060884.1:2000SEP08	4897818H1	591	876
64	LG:1060884.1:2000SEP08	4505832H1	627	876
64	LG:1060884.1:2000SEP08	1662074H1	757	866
64	LG:1060884.1:2000SEP08	6157363H1	833	931
65	LG:242191.1:2000SEP08	1670394H1	1156	1377
65	LG:242191.1:2000SEP08	1670394F6	1156	1402
65	LG:242191.1:2000SEP08	g1523196	1260	1387
65	LG:242191.1:2000SEP08	g3741377	1323	1651
65	LG:242191.1:2000SEP08	1887885F6	1045	1185
65	LG:242191.1:2000SEP08	5273771H1	1060	1215
65	LG:242191.1:2000SEP08	5273992H1	1060	1221
65	LG:242191.1:2000SEP08	612957H1	1039	1153
65	LG:242191.1:2000SEP08	5614413H1	1024	1213
65	LG:242191.1:2000SEP08	612957F1	1032	1658

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
65	LG:242191.1:2000SEP08	612957R1	1039	1528
65	LG:242191.1:2000SEP08	3119047H1	829	1105
65	LG:242191.1:2000SEP08	6339308H1	931	1051
65	LG:242191.1:2000SEP08	863189H1	965	1217
65	LG:242191.1:2000SEP08	6317077H1	967	1180
65	LG:242191.1:2000SEP08	2758331H1	1008	1277
65	LG:242191.1:2000SEP08	6307074H1	1017	1190
65	LG:242191.1:2000SEP08	g2805691	765	1165
65	LG:242191.1:2000SEP08	6985383H1	774	992
65	LG:242191.1:2000SEP08	g4763773	316	748
65	LG:242191.1:2000SEP08	g5744234	337	773
65	LG:242191.1:2000SEP08	g4523792	388	772
65	LG:242191.1:2000SEP08	879149H1	399	525
65	LG:242191.1:2000SEP08	6816123H1	497	1117
65	LG:242191.1:2000SEP08	4591471H1	1	248
65	LG:242191.1:2000SEP08	5378162H1	100	352
65	LG:242191.1:2000SEP08	g6700490	234	771
65	LG:242191.1:2000SEP08	7754528H1	237	764
65	LG:242191.1:2000SEP08	5378162F8	242	352
65	LG:242191.1:2000SEP08	6752934J1	264	763
66	LG:1063762.3:2000SEP08	g2046867	368	720
66	LG:1063762.3:2000SEP08	4140846F9	453	1035
66	LG:1063762.3:2000SEP08	6549845F8	1	688
66	LG:1063762.3:2000SEP08	1528228H1	164	367
66	LG:1063762.3:2000SEP08	1262724H1	175	391
66	LG:1063762.3:2000SEP08	7273596H1	269	887
66	LG:1063762.3:2000SEP08	2495993H1	952	1192
66	LG:1063762.3:2000SEP08	4408341H1	1055	1175
66	LG:1063762.3:2000SEP08	111294R6	899	1321
66	LG:1063762.3:2000SEP08	111294R1	899	1350
66	LG:1063762.3:2000SEP08	4140846H1	454	735
66	LG:1063762.3:2000SEP08	4030602F8	556	1087
66	LG:1063762.3:2000SEP08	3929925H1	630	905
67	LG:1100856.1:2000SEP08	6790835F8	1	545
67	LG:1100856.1:2000SEP08	6790835T8	1	455
67	LG:1100856.1:2000SEP08	6790835H1	1	276
68	LG:979390.2:2000SEP08	g922268	84	327
68	LG:979390.2:2000SEP08	g922215	84	388
68	LG:979390.2:2000SEP08	7432338H1	805	1107
68	LG:979390.2:2000SEP08	7405841H1	458	983
68	LG:979390.2:2000SEP08	4991883H1	845	903
68	LG:979390.2:2000SEP08	1958936H1	678	856
68	LG:979390.2:2000SEP08	4991883T6	268	807
68	LG:979390.2:2000SEP08	7651737J1	243	739
68	LG:979390.2:2000SEP08	g991321	109	487
68	LG:979390.2:2000SEP08	6927079H1	1	400
68	LG:979390.2:2000SEP08	4991883F6	1077	1282
68	LG:979390.2:2000SEP08	7651737H1	490	1125
68	LG:979390.2:2000SEP08	g922269	1131	1411



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
68	LG:979390.2:2000SEP08	g991279	1139	1387
68	LG:979390.2:2000SEP08	g922216	1024	1410
68	LG:979390.2:2000SEP08	g989914	1153	1399
69	LG:1400447.1:2000SEP08	1375814H1	261	292
69	LG:1400447.1:2000SEP08	1379688F6	261	374
69	LG:1400447.1:2000SEP08	7373664H1	1	396
69	LG:1400447.1:2000SEP08	1375814F1	53	431
69	LG:1400447.1:2000SEP08	1379688T6	96	629
69	LG:1400447.1:2000SEP08	g697373	422	636
70	LG:1400562.1:2000SEP08	734146R6	50	463
70	LG:1400562.1:2000SEP08	920342T6	50	463
70	LG:1400562.1:2000SEP08	913717T6	50	463
70	LG:1400562.1:2000SEP08	870924R6	50	422
70	LG:1400562.1:2000SEP08	746315R6	50	415
70	LG:1400562.1:2000SEP08	840754T6	83	463
70	LG:1400562.1:2000SEP08	871602R6	50	364
70	LG:1400562.1:2000SEP08	734229R6	50	460
70	LG:1400562.1:2000SEP08	861612T6	50	463
70	LG:1400562.1:2000SEP08	751054R6	50	463
70	LG:1400562.1:2000SEP08	914213R6	50	364
70	LG:1400562.1:2000SEP08	736850T6	77	463
70	LG:1400562.1:2000SEP08	3637162H1	38	329
70	LG:1400562.1:2000SEP08	919277R6	50	419
70	LG:1400562.1:2000SEP08	905555R6	50	419
70	LG:1400562.1:2000SEP08	913717R6	50	355
70	LG:1400562.1:2000SEP08	4162946H1	18	262
70	LG:1400562.1:2000SEP08	849924R6	50	463
70	LG:1400562.1:2000SEP08	5490338H1	1	274
70	LG:1400562.1:2000SEP08	6937225H1	35	536
70	LG:1400562.1:2000SEP08	3637162F6	38	490
70	LG:1400562.1:2000SEP08	g1696681	40	420
70	LG:1400562.1:2000SEP08	3637162T6	192	740
70	LG:1400562.1:2000SEP08	768130R6	50	463
70	LG:1400562.1:2000SEP08	833958R6	50	450
70	LG:1400562.1:2000SEP08	871602T6	50	463
70	LG:1400562.1:2000SEP08	746815R6	50	463
70	LG:1400562.1:2000SEP08	746908R6	50	364
70	LG:1400562.1:2000SEP08	840754R6	50	449
70	LG:1400562.1:2000SEP08	1708790H1	63	264
70	LG:1400562.1:2000SEP08	920553T6	50	463
70	LG:1400562.1:2000SEP08	4162946F6	18	381
70	LG:1400562.1:2000SEP08	900986T6	50	463
70	LG:1400562.1:2000SEP08	900986R6	50	364
70	LG:1400562.1:2000SEP08	919277T6	50	463
70	LG:1400562.1:2000SEP08	909272T6	50	463
70	LG:1400562.1:2000SEP08	808626R6	50	463
70	LG:1400562.1:2000SEP08	754651R6	50	364
70	LG:1400562.1:2000SEP08	905555T6	50	463
70	LG:1400562.1:2000SEP08	920553R6	50	447

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
70	LG:1400562.1:2000SEP08	849924T6	50	463
70	LG:1400562.1:2000SEP08	864660R6	50	463
70	LG:1400562.1:2000SEP08	6206095H1	28	381
70	LG:1400562.1:2000SEP08	900841T6	50	463
70	LG:1400562.1:2000SEP08	905554T6	50	463
70	LG:1400562.1:2000SEP08	745305R6	50	463
70	LG:1400562.1:2000SEP08	743895R6	50	463
70	LG:1400562.1:2000SEP08	909272R6	50	449
70	LG:1400562.1:2000SEP08	864660T6	50	463
70	LG:1400562.1:2000SEP08	734146H1	50	279
70	LG:1400562.1:2000SEP08	920342R6	50	364
70	LG:1400562.1:2000SEP08	905554R6	50	364
70	LG:1400562.1:2000SEP08	914213T6	50	463
70	LG:1400562.1:2000SEP08	736850R6	50	422
70	LG:1400562.1:2000SEP08	6200892T8	50	315
70	LG:1400562.1:2000SEP08	870924T6	50	463
70	LG:1400562.1:2000SEP08	861612R6	50	463
71	LG:1076130.1:2000SEP08	6311083F8	168	544
71	LG:1076130.1:2000SEP08	g6076180	476	810
71	LG:1076130.1:2000SEP08	3042046T6	258	768
71	LG:1076130.1:2000SEP08	5543192T6	228	767
71	LG:1076130.1:2000SEP08	6311083H1	168	725
71	LG:1076130.1:2000SEP08	g1261341	547	722
71	LG:1076130.1:2000SEP08	6936847R8	1	630
72	LG:1064459.1:2000SEP08	3128415F6	1	253
72	LG:1064459.1:2000SEP08	7155878H1	50	571
72	LG:1064459.1:2000SEP08	7278205H1	105	645
72	LG:1064459.1:2000SEP08	6176589F8	269	918
72	LG:1064459.1:2000SEP08	3201201T6	276	798
72	LG:1064459.1:2000SEP08	6880021J1	389	1021
72	LG:1064459.1:2000SEP08	2878289F6	593	1098
72	LG:1064459.1:2000SEP08	g2180031	602	957
72	LG:1064459.1:2000SEP08	2870408H1	593	869
72	LG:1064459.1:2000SEP08	2870408F7	593	1030
72	LG:1064459.1:2000SEP08	2638526H1	460	579
72	LG:1064459.1:2000SEP08	g1059584	630	901
72	LG:1064459.1:2000SEP08	2870424H1	596	806
72	LG:1064459.1:2000SEP08	2878289H1	593	873
73	LG:1079415.14:2000SEP08	7358860H1	1	370
73	LG:1079415.14:2000SEP08	g2166139	68	561
74	LG:1329431.3:2000SEP08	g2834309	46	465
74	LG:1329431.3:2000SEP08	3281151H1	328	597
74	LG:1329431.3:2000SEP08	3281151F6	327	763
74	LG:1329431.3:2000SEP08	7678170H1	1	603
74	LG:1329431.3:2000SEP08	6480874H1	297	763
75	LG:1088431.2:2000SEP08	893591H1	25	304
75	LG:1088431.2:2000SEP08	6593962H1	1	161
75	LG:1088431.2:2000SEP08	6593962F8	1	161
75	LG:1088431.2:2000SEP08	894136H1	24	191

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
75	LG:1088431.2:2000SEP08	3685359H1	45	347
75	LG:1088431.2:2000SEP08	3685359F6	45	504
75	LG:1088431.2:2000SEP08	3685359T6	267	545
76	LG:1329462.2:2000SEP08	1567760H1	1011	1230
76	LG:1329462.2:2000SEP08	7430912H1	986	1295
76	LG:1329462.2:2000SEP08	6732779H1	807	985
76	LG:1329462.2:2000SEP08	6844540T8	1099	1277
76	LG:1329462.2:2000SEP08	6537070H1	18	539
76	LG:1329462.2:2000SEP08	g2987926	1043	1235
76	LG:1329462.2:2000SEP08	7007661H1	18	482
76	LG:1329462.2:2000SEP08	7218223H1	1	510
76	LG:1329462.2:2000SEP08	7437976H1	54	324
76	LG:1329462.2:2000SEP08	7634846J1	417	973
76	LG:1329462.2:2000SEP08	5200862F6	573	1082
76	LG:1329462.2:2000SEP08	4630864F8	624	870
76	LG:1329462.2:2000SEP08	6771021J1	733	1285
76	LG:1329462.2:2000SEP08	5387945T6	1004	1636
76	LG:1329462.2:2000SEP08	1567764F6	1011	1349
76	LG:1329462.2:2000SEP08	6939863R8	1166	1547
76	LG:1329462.2:2000SEP08	g4438778	1218	1603
76	LG:1329462.2:2000SEP08	4043836F6	1353	1596
76	LG:1329462.2:2000SEP08	4043836H1	1354	1593
76	LG:1329462.2:2000SEP08	4043836F9	1359	1596
77	LI:393468.1:2000SEP08	7619106J1	444	979
77	LI:393468.1:2000SEP08	g6040695	217	669
77	LI:393468.1:2000SEP08	g6836681	222	664
77	LI:393468.1:2000SEP08	g3918185	266	664
77	LI:393468.1:2000SEP08	g4525302	265	664
77	LI:393468.1:2000SEP08	g6991355	260	663
77	LI:393468.1:2000SEP08	8023621J1	47	655
77	LI:393468.1:2000SEP08	g4124201	284	646
77	LI:393468.1:2000SEP08	g4371399	169	646
77	LI:393468.1:2000SEP08	g3144307	196	644
77	LI:393468.1:2000SEP08	1657896F6	1	245
77	LI:393468.1:2000SEP08	1657896H1	1	234
78	LI:722577.1:2000SEP08	6270214F8	1	591
78	LI:722577.1:2000SEP08	6270214H2	1	514
78	LI:722577.1:2000SEP08	6270214T8	33	484
79	LI:322783.16:2000SEP08	6461340H1	1	538
79	LI:322783.16:2000SEP08	6461140H2	2	393
79	LI:322783.16:2000SEP08	g1259601	8	227
79	LI:322783.16:2000SEP08	g1988731	38	314
79	LI:322783.16:2000SEP08	6526629H1	170	703
79	LI:322783.16:2000SEP08	3483529H1	266	590
79	LI:322783.16:2000SEP08	g705596	332	672
79	LI:322783.16:2000SEP08	5703243H1	382	663
79	LI:322783.16:2000SEP08	7604365H1	311	688
80	LI:901355.2:2000SEP08	4664370F6	1	571
80	LI:901355.2:2000SEP08	8004786H1	34	554

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
81	LI:038859.2:2000SEP08	8044253H1	325	976
81	LI:038859.2:2000SEP08	g6699872	365	907
81	LI:038859.2:2000SEP08	g6703559	525	907
81	LI:038859.2:2000SEP08	7386340H1	577	899
81	LI:038859.2:2000SEP08	6763732J1	311	898
81	LI:038859.2:2000SEP08	g7153702	470	875
81	LI:038859.2:2000SEP08	g5177300	422	866
81	LI:038859.2:2000SEP08	g3041438	688	858
81	LI:038859.2:2000SEP08	g2737221	467	847
81	LI:038859.2:2000SEP08	g2218499	631	843
81	LI:038859.2:2000SEP08	g2574686	551	828
81	LI:038859.2:2000SEP08	6777292H1	214	828
81	LI:038859.2:2000SEP08	6777292J1	1	650
81	LI:038859.2:2000SEP08	6986658H1	207	307
82	LI:1046117.1:2000SEP08	g4740070	79	305
82	LI:1046117.1:2000SEP08	8100423H1	1	543
83	LI:801015.1:2000SEP08	2007652H1	1	187
83	LI:801015.1:2000SEP08	2007652T6	1	225
83	LI:801015.1:2000SEP08	2007652R6	1	256
84	LI:1175590.1:2000SEP08	3288387T6	329	715
84	LI:1175590.1:2000SEP08	4219127T8	326	715
84	LI:1175590.1:2000SEP08	4219127T6	516	715
84	LI:1175590.1:2000SEP08	3288387F6	1	420
84	LI:1175590.1:2000SEP08	3288387H1	1	251
85	LI:1170585.2:2000SEP08	618016H1	1	282
85	LI:1170585.2:2000SEP08	618016R6	1	258
85	LI:1170585.2:2000SEP08	2501520H1	18	249
85	LI:1170585.2:2000SEP08	1359580H1	47	291
85	LI:1170585.2:2000SEP08	1359580F6	47	442
85	LI:1170585.2:2000SEP08	3012558H1	49	184
85	LI:1170585.2:2000SEP08	3021092H1	51	317
85	LI:1170585.2:2000SEP08	7579667H1	54	330
85	LI:1170585.2:2000SEP08	6206808H1	115	526
85	LI:1170585.2:2000SEP08	5502284H1	133	356
85	LI:1170585.2:2000SEP08	2498102H1	137	296
85	LI:1170585.2:2000SEP08	5495221H1	243	487
85	LI:1170585.2:2000SEP08	618016T6	288	726
85	LI:1170585.2:2000SEP08	g3238505	374	788
85	LI:1170585.2:2000SEP08	4029953H1	416	670
86	LI:719531.2:2000SEP08	7739059J1	1	207
87	LI:794623.1:2000SEP08	3209591H1	39	184
87	LI:794623.1:2000SEP08	5642563H1	1	245
87	LI:794623.1:2000SEP08	g3190333	43	117
88	LI:1173119.1:2000SEP08	4896201F8	1	427
88	LI:1173119.1:2000SEP08	4896201H1	1	269
88	LI:1173119.1:2000SEP08	2717909H1	132	387
88	LI:1173119.1:2000SEP08	392443R1	288	816
88	LI:1173119.1:2000SEP08	g3038179	679	778
88	LI:1173119.1:2000SEP08	3989646R6	57	451

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
88	LI:1173119.1:2000SEP08	828479H1	320	577
88	LI:1173119.1:2000SEP08	2642591H1	350	587
88	LI:1173119.1:2000SEP08	3989646H1	57	215
88	LI:1173119.1:2000SEP08	7933083H1	92	601
89	LI:1093285.1:2000SEP08	7665924H1	711	769
89	LI:1093285.1:2000SEP08	6934423R9	1	630
89	LI:1093285.1:2000SEP08	6781646H1	300	778
89	LI:1093285.1:2000SEP08	6781646J1	357	778
89	LI:1093285.1:2000SEP08	4524026H1	473	734
89	LI:1093285.1:2000SEP08	6781646F8	300	778
89	LI:1093285.1:2000SEP08	6254513T8	568	770
89	LI:1093285.1:2000SEP08	6781646R8	528	770
89	LI:1093285.1:2000SEP08	2591085H2	387	541
90	LI:1091881.1:2000SEP08	4531208H1	121	405
90	LI:1091881.1:2000SEP08	4982821H1	1	269
90	LI:1091881.1:2000SEP08	4982821F7	1	476
90	LI:1091881.1:2000SEP08	6140274H1	16	324
90	LI:1091881.1:2000SEP08	g2618444	390	476
90	LI:1091881.1:2000SEP08	g6714003	227	476
91	LI:1091617.1:2000SEP08	4642113H1	1	273
91	LI:1091617.1:2000SEP08	7060192H1	203	784
91	LI:1091617.1:2000SEP08	7060192T8	370	913
91	LI:1091617.1:2000SEP08	7060192F8	216	364
92	LI:1082344.1:2000SEP08	1706514H1	1	226
92	LI:1082344.1:2000SEP08	1706514F6	1	411
92	LI:1082344.1:2000SEP08	6002889H1	166	426
92	LI:1082344.1:2000SEP08	6002889F8	166	827
92	LI:1082344.1:2000SEP08	6002889T8	270	779
93	LI:1166249.1:2000SEP08	2817644H1	40	351
93	LI:1166249.1:2000SEP08	2818537H1	40	256
93	LI:1166249.1:2000SEP08	294959H1	514	594
93	LI:1166249.1:2000SEP08	2829038H1	1	231
93	LI:1166249.1:2000SEP08	g1635160	1	146
93	LI:1166249.1:2000SEP08	438226H1	5	148
93	LI:1166249.1:2000SEP08	2818537F6	40	594
94	LI:799675.1:2000SEP08	1208995T6	83	133
94	LI:799675.1:2000SEP08	1208995H1	1	128
94	LI:799675.1:2000SEP08	4770882H1	37	189
94	LI:799675.1:2000SEP08	2498778T6	56	189
94	LI:799675.1:2000SEP08	4876022F6	74	501
94	LI:799675.1:2000SEP08	1208995R6	83	133
94	LI:799675.1:2000SEP08	7988786H1	83	133
94	LI:799675.1:2000SEP08	4876022H1	75	340
95	LI:1178899.1:2000SEP08	5387945T6	1023	1655
95	LI:1178899.1:2000SEP08	g4438778	1237	1622
95	LI:1178899.1:2000SEP08	6939863R8	1185	1566
95	LI:1178899.1:2000SEP08	1567764F6	1030	1368
95	LI:1178899.1:2000SEP08	5200862F6	594	1101
95	LI:1178899.1:2000SEP08	7634846J1	438	994

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
95	LI:1178899.1:2000SEP08	5200862H1	594	813
95	LI:1178899.1:2000SEP08	6537070H1	1	560
95	LI:1178899.1:2000SEP08	1567760H1	1030	1249
96	LI:1169241.1:2000SEP08	g3785390	229	632
96	LI:1169241.1:2000SEP08	g316080	250	495
96	LI:1169241.1:2000SEP08	3079252H2	428	721
96	LI:1169241.1:2000SEP08	7585969H2	1	361
96	LI:1169241.1:2000SEP08	60113610D2	1	415
96	LI:1169241.1:2000SEP08	60120205D2	1	451
96	LI:1169241.1:2000SEP08	g4684943	542	632
96	LI:1169241.1:2000SEP08	60106944D1	1	299
96	LI:1169241.1:2000SEP08	60209330U1	322	610
96	LI:1169241.1:2000SEP08	g4328800	542	632
96	LI:1169241.1:2000SEP08	3077177F6	429	610
96	LI:1169241.1:2000SEP08	3077177H1	429	610
96	LI:1169241.1:2000SEP08	60209329U1	322	610
96	LI:1169241.1:2000SEP08	60106937D1	1	323
97	LI:1180090.1:2000SEP08	5326859T9	248	828
97	LI:1180090.1:2000SEP08	5324689T9	300	825
97	LI:1180090.1:2000SEP08	5326859F8	162	747
97	LI:1180090.1:2000SEP08	5311657F8	1	650
97	LI:1180090.1:2000SEP08	5326859H1	162	398
98	LI:2049322.1:2000SEP08	1698730H1	164	358
98	LI:2049322.1:2000SEP08	4727660H1	1	266
98	LI:2049322.1:2000SEP08	4822628H1	1	282
98	LI:2049322.1:2000SEP08	1322674H1	1	255
98	LI:2049322.1:2000SEP08	1798649F6	5	360
98	LI:2049322.1:2000SEP08	g3736671	86	515
98	LI:2049322.1:2000SEP08	1698965F6	164	528
98	LI:2049322.1:2000SEP08	1698965T6	178	499
98	LI:2049322.1:2000SEP08	1698965H1	164	395
98	LI:2049322.1:2000SEP08	1798649H1	5	279
99	LI:809074.1:2000SEP08	455966H1	20	255
99	LI:809074.1:2000SEP08	5998240F8	1	398
99	LI:809074.1:2000SEP08	5998240T8	1	408
99	LI:809074.1:2000SEP08	5998240H1	1	484
99	LI:809074.1:2000SEP08	455162R6	20	521
99	LI:809074.1:2000SEP08	3206542H1	25	202
99	LI:809074.1:2000SEP08	7652675J1	47	172
99	LI:809074.1:2000SEP08	g2401959	129	486
99	LI:809074.1:2000SEP08	460544H1	20	245
99	LI:809074.1:2000SEP08	g5101146	180	486
99	LI:809074.1:2000SEP08	458061H1	20	259
99	LI:809074.1:2000SEP08	455162H1	20	272
99	LI:809074.1:2000SEP08	460544R6	20	326
99	LI:809074.1:2000SEP08	4663964H1	37	289
99	LI:809074.1:2000SEP08	g1807165	88	278
99	LI:809074.1:2000SEP08	3614680H1	155	272
100	LI:805158.1:2000SEP08	4028040F6	1	366

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
100	LI:805158.1:2000SEP08	4028295H1	2	259
100	LI:805158.1:2000SEP08	7174811T8	213	528
100	LI:805158.1:2000SEP08	g1424215	300	756
100	LI:805158.1:2000SEP08	5284519T9	442	529
100	LI:805158.1:2000SEP08	5284519H1	442	645
100	LI:805158.1:2000SEP08	5284519F7	477	658
100	LI:805158.1:2000SEP08	5284519F8	447	643
101	LI:1172697.1:2000SEP08	4865460H1	29	276
101	LI:1172697.1:2000SEP08	7005130H1	5	566
101	LI:1172697.1:2000SEP08	7278360H1	90	585
101	LI:1172697.1:2000SEP08	6950314R8	1	444
101	LI:1172697.1:2000SEP08	7948222J1	399	892
101	LI:1172697.1:2000SEP08	7948226J1	399	892
101	LI:1172697.1:2000SEP08	7948226H1	409	892
101	LI:1172697.1:2000SEP08	7460010H1	430	570
101	LI:1172697.1:2000SEP08	5154688H1	607	848
101	LI:1172697.1:2000SEP08	7162245F8	1	513
101	LI:1172697.1:2000SEP08	7100025F8	68	630
101	LI:1172697.1:2000SEP08	g2027810	19	261
101	LI:1172697.1:2000SEP08	5071474F8	15	211
101	LI:1172697.1:2000SEP08	55027336H1	161	609
101	LI:1172697.1:2000SEP08	7100025H1	68	591
101	LI:1172697.1:2000SEP08	g2027676	19	385
101	LI:1172697.1:2000SEP08	8016160J1	486	887
101	LI:1172697.1:2000SEP08	7946870H1	1	330
101	LI:1172697.1:2000SEP08	7162245H1	1	379
101	LI:1172697.1:2000SEP08	7948222H1	410	892
101	LI:1172697.1:2000SEP08	7580707H1	2	340
101	LI:1172697.1:2000SEP08	g5706987	470	637
101	LI:1172697.1:2000SEP08	7946870J1	264	649
102	LI:1174107.2:2000SEP08	5501966H1	9	271
102	LI:1174107.2:2000SEP08	8098550H1	1	499
102	LI:1174107.2:2000SEP08	8107387J1	66	584
102	LI:1174107.2:2000SEP08	2818123T6	102	357
102	LI:1174107.2:2000SEP08	8107387H1	228	584
102	LI:1174107.2:2000SEP08	4028717H1	401	670
102	LI:1174107.2:2000SEP08	6441209H1	107	331
102	LI:1174107.2:2000SEP08	5649356H1	9	333
103	LI:1177434.2:2000SEP08	5310369F6	1	431
103	LI:1177434.2:2000SEP08	5310369F8	12	423
103	LI:1177434.2:2000SEP08	g1081517	13	367
103	LI:1177434.2:2000SEP08	g1985777	236	435
103	LI:1177434.2:2000SEP08	4265263H1	253	336
103	LI:1177434.2:2000SEP08	g1984842	306	615
103	LI:1177434.2:2000SEP08	g3739050	310	636
103	LI:1177434.2:2000SEP08	6781939J1	331	941
103	LI:1177434.2:2000SEP08	g179296	397	876
103	LI:1177434.2:2000SEP08	8104872H1	408	1042
103	LI:1177434.2:2000SEP08	8104872J1	680	1285

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
103	LI:1177434.2:2000SEP08	097477H1	844	1032
103	LI:1177434.2:2000SEP08	5310369H1	1	243
104	LI:1184255.1:2000SEP08	5261341H1	1	219
104	LI:1184255.1:2000SEP08	5261341F6	1	479
104	LI:1184255.1:2000SEP08	5953264H1	50	175
104	LI:1184255.1:2000SEP08	5953264F8	51	779
104	LI:1184255.1:2000SEP08	1852536H1	283	570
104	LI:1184255.1:2000SEP08	g4070788	336	774
104	LI:1184255.1:2000SEP08	g2806837	418	773
104	LI:1184255.1:2000SEP08	g3148317	459	771
105	LI:1164555.1:2000SEP08	2813750F6	1	471
105	LI:1164555.1:2000SEP08	2813750T6	1	487
106	LI:238666.4:2000SEP08	3929925H1	630	905
106	LI:238666.4:2000SEP08	111294R6	899	1320
106	LI:238666.4:2000SEP08	g2046867	368	720
106	LI:238666.4:2000SEP08	653171H1	392	542
106	LI:238666.4:2000SEP08	4140846F9	453	1035
106	LI:238666.4:2000SEP08	4140846H1	454	735
106	LI:238666.4:2000SEP08	4030602F8	556	1087
106	LI:238666.4:2000SEP08	111294R1	899	1320
106	LI:238666.4:2000SEP08	4030602H1	573	653
106	LI:238666.4:2000SEP08	2495993H1	952	1192
106	LI:238666.4:2000SEP08	6549845H1	1	407
106	LI:238666.4:2000SEP08	6549845F8	1	688
106	LI:238666.4:2000SEP08	7273596H1	269	887
107	LI:1166752.1:2000SEP08	5500807H1	1	180
107	LI:1166752.1:2000SEP08	5497912H1	1	260
107	LI:1166752.1:2000SEP08	1004710H1	146	384
107	LI:1166752.1:2000SEP08	2666231H1	270	523
107	LI:1166752.1:2000SEP08	5500805H1	1	271
107	LI:1166752.1:2000SEP08	2866837H1	344	654
107	LI:1166752.1:2000SEP08	5500805F8	1	581
107	LI:1166752.1:2000SEP08	2666231F6	165	523
107	LI:1166752.1:2000SEP08	5497912F9	4	439
108	LI:2049654.1:2000SEP08	7373664H1	1	344
108	LI:2049654.1:2000SEP08	1375814F1	1	379
108	LI:2049654.1:2000SEP08	1375814F6	1	361
108	LI:2049654.1:2000SEP08	7705726J1	38	586
108	LI:2049654.1:2000SEP08	1379688T6	41	577
108	LI:2049654.1:2000SEP08	g697373	370	584
108	LI:2049654.1:2000SEP08	1379688F6	1	322
108	LI:2049654.1:2000SEP08	1375814H1	1	240
108	LI:2049654.1:2000SEP08	1379688H1	1	222
109	LI:242665.2:2000SEP08	4155245F6	1	472
109	LI:242665.2:2000SEP08	4155245H1	1	249
109	LI:242665.2:2000SEP08	2815027H1	207	439
110	LI:208637.1:2000SEP08	2045901H1	3997	4170
110	LI:208637.1:2000SEP08	511147H1	4004	4308
110	LI:208637.1:2000SEP08	511147R6	4004	4294



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
110	LI:208637.1:2000SEP08	g6993309	1	450
110	LI:208637.1:2000SEP08	g5914430	9	460
110	LI:208637.1:2000SEP08	g3679677	20	483
110	LI:208637.1:2000SEP08	g4190232	20	485
110	LI:208637.1:2000SEP08	g6710036	36	491
110	LI:208637.1:2000SEP08	g4852090	39	495
110	LI:208637.1:2000SEP08	g4305494	40	505
110	LI:208637.1:2000SEP08	g5178218	79	500
110	LI:208637.1:2000SEP08	7355332H1	277	856
110	LI:208637.1:2000SEP08	7723481J2	328	846
110	LI:208637.1:2000SEP08	1551301R6	3397	3627
110	LI:208637.1:2000SEP08	1551301H1	3397	3542
110	LI:208637.1:2000SEP08	g6570747	3561	3805
110	LI:208637.1:2000SEP08	3415286H1	4263	4530
110	LI:208637.1:2000SEP08	70455975V1	4291	4710
110	LI:208637.1:2000SEP08	g787701	4288	4618
110	LI:208637.1:2000SEP08	446023H1	4004	4246
110	LI:208637.1:2000SEP08	442886H1	4004	4246
110	LI:208637.1:2000SEP08	2155512H1	4520	4770
110	LI:208637.1:2000SEP08	g2583420	4529	4941
110	LI:208637.1:2000SEP08	g2155854	4530	5030
110	LI:208637.1:2000SEP08	3417104H1	4531	4765
110	LI:208637.1:2000SEP08	g787685	4288	4530
110	LI:208637.1:2000SEP08	2445053F6	4291	4804
110	LI:208637.1:2000SEP08	2445053H1	4291	4526
110	LI:208637.1:2000SEP08	1662789H1	4296	4528
110	LI:208637.1:2000SEP08	5217068H1	4298	4560
110	LI:208637.1:2000SEP08	1663583F6	4296	4628
110	LI:208637.1:2000SEP08	4531059H1	3128	3402
110	LI:208637.1:2000SEP08	7081386H1	3147	3663
110	LI:208637.1:2000SEP08	5394846H1	3949	4250
110	LI:208637.1:2000SEP08	4536525H1	3962	4215
110	LI:208637.1:2000SEP08	7598517H1	3989	4561
110	LI:208637.1:2000SEP08	5394325H1	3992	4281
110	LI:208637.1:2000SEP08	042108H1	3999	4207
110	LI:208637.1:2000SEP08	6991133H1	753	1021
110	LI:208637.1:2000SEP08	8039312J1	930	1424
110	LI:208637.1:2000SEP08	7617296J1	946	1442
110	LI:208637.1:2000SEP08	8039151J1	1454	2144
110	LI:208637.1:2000SEP08	8040751H1	1209	1852
110	LI:208637.1:2000SEP08	6623125H1	1930	2419
110	LI:208637.1:2000SEP08	6819912F8	1622	2159
110	LI:208637.1:2000SEP08	6896576H1	1625	2117
110	LI:208637.1:2000SEP08	8039312H1	1707	2308
110	LI:208637.1:2000SEP08	7984277H1	1888	2495
110	LI:208637.1:2000SEP08	6196362H1	3870	4234
110	LI:208637.1:2000SEP08	4880402H1	3874	4140
110	LI:208637.1:2000SEP08	3468193H1	3894	4139
110	LI:208637.1:2000SEP08	365246R6	3906	4235

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
110	LI:208637.1:2000SEP08	g2079127	3910	4299
110	LI:208637.1:2000SEP08	6480435H1	2338	2879
110	LI:208637.1:2000SEP08	6425484H1	2026	2536
110	LI:208637.1:2000SEP08	7617296H1	2303	2945
110	LI:208637.1:2000SEP08	2555627F6	2024	2590
110	LI:208637.1:2000SEP08	2555627H1	2025	2270
110	LI:208637.1:2000SEP08	6425484F8	2026	2547
110	LI:208637.1:2000SEP08	5924864H1	4755	5056
110	LI:208637.1:2000SEP08	860218H1	4757	5000
110	LI:208637.1:2000SEP08	323803H1	4764	5065
110	LI:208637.1:2000SEP08	4126367H1	4764	5039
110	LI:208637.1:2000SEP08	70453526V1	4771	5224
110	LI:208637.1:2000SEP08	70453430V1	4772	5203
110	LI:208637.1:2000SEP08	2721776H1	4637	4887
110	LI:208637.1:2000SEP08	g2209684	4651	4944
110	LI:208637.1:2000SEP08	1500756T6	4662	5165
110	LI:208637.1:2000SEP08	g2261582	4671	4944
110	LI:208637.1:2000SEP08	3408148H1	4722	4849
110	LI:208637.1:2000SEP08	2873391H1	4726	5009
110	LI:208637.1:2000SEP08	4763672H1	4728	4949
110	LI:208637.1:2000SEP08	2555275H1	4735	4997
110	LI:208637.1:2000SEP08	g2912948	4746	5181
110	LI:208637.1:2000SEP08	6783219H1	4752	5202
110	LI:208637.1:2000SEP08	g2583727	4774	5202
110	LI:208637.1:2000SEP08	g4283835	4785	5204
110	LI:208637.1:2000SEP08	g4437576	4787	5203
110	LI:208637.1:2000SEP08	g2583781	4788	5213
110	LI:208637.1:2000SEP08	g4874731	4794	5209
110	LI:208637.1:2000SEP08	g3094261	4797	5204
110	LI:208637.1:2000SEP08	7338219H1	4801	5207
110	LI:208637.1:2000SEP08	g2576977	4800	5213
110	LI:208637.1:2000SEP08	7337919H1	4800	5203
110	LI:208637.1:2000SEP08	511147T6	4805	5164
110	LI:208637.1:2000SEP08	g4112440	4822	5206
110	LI:208637.1:2000SEP08	365246T6	4823	5166
110	LI:208637.1:2000SEP08	g1920661	4826	5203
110	LI:208637.1:2000SEP08	g2188517	4835	5203
110	LI:208637.1:2000SEP08	g1941487	4875	5211
110	LI:208637.1:2000SEP08	4506729H1	4875	5150
110	LI:208637.1:2000SEP08	g787649	4882	5180
110	LI:208637.1:2000SEP08	g3898699	4907	5203
110	LI:208637.1:2000SEP08	1663583T6	4919	5163
110	LI:208637.1:2000SEP08	4144853H1	4937	5203
110	LI:208637.1:2000SEP08	2607654T6	4966	5158
110	LI:208637.1:2000SEP08	2607654F6	4971	5203
110	LI:208637.1:2000SEP08	2607654H1	4971	5203
110	LI:208637.1:2000SEP08	3470176H1	4976	5203
110	LI:208637.1:2000SEP08	g787635	4983	5203
110	LI:208637.1:2000SEP08	g7278523	5067	5203

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
110	LI:208637.1:2000SEP08	g664248	5113	5204
110	LI:208637.1:2000SEP08	1680896H1	5116	5205
110	LI:208637.1:2000SEP08	2419545H1	5153	5203
110	LI:208637.1:2000SEP08	4277393H1	3811	3936
110	LI:208637.1:2000SEP08	7617468H1	498	1093
110	LI:208637.1:2000SEP08	5664828H1	3835	4123
110	LI:208637.1:2000SEP08	3101689H1	728	996
110	LI:208637.1:2000SEP08	3101689F6	740	1055
110	LI:208637.1:2000SEP08	6991888H1	532	1017
110	LI:208637.1:2000SEP08	g2007169	2014	2311
110	LI:208637.1:2000SEP08	2851540H1	4017	4297
110	LI:208637.1:2000SEP08	2723689H1	3827	4067
110	LI:208637.1:2000SEP08	3270687H1	4024	4267
110	LI:208637.1:2000SEP08	2287401H1	4059	4328
110	LI:208637.1:2000SEP08	2968020H1	4104	4400
110	LI:208637.1:2000SEP08	2536769H1	4129	4374
110	LI:208637.1:2000SEP08	6517121H1	4161	4495
110	LI:208637.1:2000SEP08	2440915H1	4154	4402
110	LI:208637.1:2000SEP08	6480435F9	2353	2948
110	LI:208637.1:2000SEP08	2730768H1	2435	2677
110	LI:208637.1:2000SEP08	4071601H1	2650	2963
110	LI:208637.1:2000SEP08	4071601F8	2649	3200
110	LI:208637.1:2000SEP08	7251820F8	2682	3228
110	LI:208637.1:2000SEP08	g1941486	4250	4448
110	LI:208637.1:2000SEP08	3293726H1	4547	4819
110	LI:208637.1:2000SEP08	6441190H1	4544	4790
110	LI:208637.1:2000SEP08	5025609H1	4548	4800
110	LI:208637.1:2000SEP08	2876024H1	4580	4874
110	LI:208637.1:2000SEP08	3799471H1	4592	4878
110	LI:208637.1:2000SEP08	2445053T6	4634	5165
110	LI:208637.1:2000SEP08	g4002423	4633	5088
110	LI:208637.1:2000SEP08	g4071919	4637	5085
110	LI:208637.1:2000SEP08	g2013196	3733	4051
110	LI:208637.1:2000SEP08	g1920720	3752	4136
110	LI:208637.1:2000SEP08	5399720H1	3259	3514
110	LI:208637.1:2000SEP08	3295282H1	3255	3504
110	LI:208637.1:2000SEP08	2777229H1	4401	4658
110	LI:208637.1:2000SEP08	3720127H1	4469	4781
110	LI:208637.1:2000SEP08	g2270753	4489	4939
110	LI:208637.1:2000SEP08	1366744R1	4408	4815
110	LI:208637.1:2000SEP08	2607995H1	4407	4669
110	LI:208637.1:2000SEP08	1366744H1	4408	4707
110	LI:208637.1:2000SEP08	1366966H1	4408	4667
110	LI:208637.1:2000SEP08	6896713H1	1588	1731
110	LI:208637.1:2000SEP08	6896713F8	1616	2181
110	LI:208637.1:2000SEP08	6896576F8	1616	2097
110	LI:208637.1:2000SEP08	7251820H1	2689	3224
110	LI:208637.1:2000SEP08	5725740H1	3821	4357
110	LI:208637.1:2000SEP08	1267290F1	4348	4626

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
110	LI:208637.1:2000SEP08	1663583H1	4296	4540
110	LI:208637.1:2000SEP08	2123867H1	4321	4603
110	LI:208637.1:2000SEP08	4934387H1	4339	4628
110	LI:208637.1:2000SEP08	1267290H1	4348	4607
110	LI:208637.1:2000SEP08	1267296H1	4348	4605
110	LI:208637.1:2000SEP08	868571H1	4348	4546
110	LI:208637.1:2000SEP08	1563256H1	4390	4624
110	LI:208637.1:2000SEP08	2151488H1	4390	4520
110	LI:208637.1:2000SEP08	2759784H1	4330	4619
110	LI:208637.1:2000SEP08	4934587H1	4330	4474
110	LI:208637.1:2000SEP08	1500756F6	3705	4143
110	LI:208637.1:2000SEP08	7578134H2	1967	2081
110	LI:208637.1:2000SEP08	1500756H1	3705	3884
110	LI:208637.1:2000SEP08	7926048H1	3813	4418
110	LI:208637.1:2000SEP08	4071601F9	2739	2831
110	LI:208637.1:2000SEP08	7079067H1	2888	3305
110	LI:208637.1:2000SEP08	3521710H1	2986	3327
110	LI:208637.1:2000SEP08	6533521F8	4161	4595
110	LI:208637.1:2000SEP08	4274884H1	4162	4408
110	LI:208637.1:2000SEP08	4220546H1	4204	4290
110	LI:208637.1:2000SEP08	5677802H1	4238	4470
110	LI:208637.1:2000SEP08	6315311H1	4241	4816
110	LI:208637.1:2000SEP08	1391430H1	4244	4501
111	LI:2051808.1:2000SEP08	6796144F8	1	553
111	LI:2051808.1:2000SEP08	6796513F8	1	541
111	LI:2051808.1:2000SEP08	6796513H1	1	454
111	LI:2051808.1:2000SEP08	6796144H1	4	444
111	LI:2051808.1:2000SEP08	6796144T8	423	998
111	LI:2051808.1:2000SEP08	6796513T8	467	993
112	LI:1175136.1:2000SEP08	70290980V1	1	447
112	LI:1175136.1:2000SEP08	70291065V1	1	447
112	LI:1175136.1:2000SEP08	70290994V1	1	205
112	LI:1175136.1:2000SEP08	70290950V1	1	447
112	LI:1175136.1:2000SEP08	70290989V1	1	468
112	LI:1175136.1:2000SEP08	70558078V1	2	223
112	LI:1175136.1:2000SEP08	70291090V1	1	388
112	LI:1175136.1:2000SEP08	70557993V1	1	260
112	LI:1175136.1:2000SEP08	70558118V1	1	207
112	LI:1175136.1:2000SEP08	70290953V1	1	457
112	LI:1175136.1:2000SEP08	5347358H1	2	278
112	LI:1175136.1:2000SEP08	70558214V1	1	158
112	LI:1175136.1:2000SEP08	70290932V1	14	488
112	LI:1175136.1:2000SEP08	70558013V1	76	360
112	LI:1175136.1:2000SEP08	70290973V1	184	421
112	LI:1175136.1:2000SEP08	70291061V1	184	407
112	LI:1175136.1:2000SEP08	70558064V1	385	447
112	LI:1175136.1:2000SEP08	70290943V1	1	470
112	LI:1175136.1:2000SEP08	70290975V1	1	470
112	LI:1175136.1:2000SEP08	70291077V1	1	451

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
112	LI:1175136.1:2000SEP08	70290908V1	1	470
113	LI:1177337.1:2000SEP08	7199154H1	1	235
113	LI:1177337.1:2000SEP08	60221095D1	1016	1454
113	LI:1177337.1:2000SEP08	7101378R8	1027	1547
113	LI:1177337.1:2000SEP08	60221094D1	1106	1492
113	LI:1177337.1:2000SEP08	60221092D1	1114	1492
113	LI:1177337.1:2000SEP08	g7376590	1422	1843
113	LI:1177337.1:2000SEP08	6256765F8	1	527
113	LI:1177337.1:2000SEP08	60221089D1	257	580
113	LI:1177337.1:2000SEP08	7101378F8	415	1047
113	LI:1177337.1:2000SEP08	2397546T6	813	1206
113	LI:1177337.1:2000SEP08	60221091D1	856	1003
113	LI:1177337.1:2000SEP08	60221093D1	996	1461
113	LI:1177337.1:2000SEP08	71256912V1	1	378
114	LI:1165056.1:2000SEP08	2728986F6	1	576
114	LI:1165056.1:2000SEP08	2728986H1	1	245
114	LI:1165056.1:2000SEP08	6476543H1	3	222
114	LI:1165056.1:2000SEP08	4613993F8	30	583
114	LI:1165056.1:2000SEP08	4613993H1	30	232
114	LI:1165056.1:2000SEP08	g989787	42	311
114	LI:1165056.1:2000SEP08	6779728J1	236	824
114	LI:1165056.1:2000SEP08	g3694145	480	862
114	LI:1165056.1:2000SEP08	6245903F8	574	1001
114	LI:1165056.1:2000SEP08	6245903H1	574	986
114	LI:1165056.1:2000SEP08	g4333202	594	861
114	LI:1165056.1:2000SEP08	7695736H1	827	1434
114	LI:1165056.1:2000SEP08	7088575H1	1101	1495
114	LI:1165056.1:2000SEP08	7284751H1	1227	1490
115	LI:1175250.1:2000SEP08	70412980D1	638	1148
115	LI:1175250.1:2000SEP08	70410311D1	348	637
115	LI:1175250.1:2000SEP08	70412242D1	79	520
115	LI:1175250.1:2000SEP08	70412921D1	1063	1239
115	LI:1175250.1:2000SEP08	70412953D1	207	637
115	LI:1175250.1:2000SEP08	70411329D1	924	1220
115	LI:1175250.1:2000SEP08	70411900D1	521	996
115	LI:1175250.1:2000SEP08	70412300D1	180	790
115	LI:1175250.1:2000SEP08	70413769D1	522	1000
115	LI:1175250.1:2000SEP08	70413678D1	521	806
115	LI:1175250.1:2000SEP08	70410754D1	665	1240
115	LI:1175250.1:2000SEP08	70413294D1	521	901
115	LI:1175250.1:2000SEP08	70413914D1	522	931
115	LI:1175250.1:2000SEP08	70412947D1	521	998
115	LI:1175250.1:2000SEP08	70410723D1	715	1238
115	LI:1175250.1:2000SEP08	70411320D1	82	520
115	LI:1175250.1:2000SEP08	70411913D1	68	491
115	LI:1175250.1:2000SEP08	70411978D1	522	1080
115	LI:1175250.1:2000SEP08	70411984D1	708	1213
115	LI:1175250.1:2000SEP08	70414092D1	182	500
115	LI:1175250.1:2000SEP08	70412562D1	144	520

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
115	U:1175250.1:2000SEP08	70413754D1	210	634
115	U:1175250.1:2000SEP08	70413246D1	678	1221
115	U:1175250.1:2000SEP08	70411843D1	179	627
115	U:1175250.1:2000SEP08	70411133D1	341	520
115	U:1175250.1:2000SEP08	70412521D1	81	520
115	U:1175250.1:2000SEP08	70413370D1	206	637
115	U:1175250.1:2000SEP08	70412332D1	405	916
115	U:1175250.1:2000SEP08	70412038D1	731	1238
115	U:1175250.1:2000SEP08	70413123D1	181	635
115	U:1175250.1:2000SEP08	70234250V1	1162	1239
115	U:1175250.1:2000SEP08	70234253V1	1164	1239
115	U:1175250.1:2000SEP08	344179T6	1170	1239
115	U:1175250.1:2000SEP08	70234444V1	1086	1220
115	U:1175250.1:2000SEP08	70234190V1	1147	1221
115	U:1175250.1:2000SEP08	70233834V1	1	183
115	U:1175250.1:2000SEP08	70411063D1	1	508
115	U:1175250.1:2000SEP08	70412905D1	48	520
115	U:1175250.1:2000SEP08	70410609D1	180	637
115	U:1175250.1:2000SEP08	70411335D1	272	863
115	U:1175250.1:2000SEP08	70234653V1	373	636
115	U:1175250.1:2000SEP08	70233964V1	382	630
115	U:1175250.1:2000SEP08	70234933V1	429	677
115	U:1175250.1:2000SEP08	7652675H1	521	1099
115	U:1175250.1:2000SEP08	70234815V1	539	826
115	U:1175250.1:2000SEP08	70234597V1	543	965
115	U:1175250.1:2000SEP08	70413945D1	639	1157
115	U:1175250.1:2000SEP08	70412845D1	640	1252
115	U:1175250.1:2000SEP08	70412776D1	674	1240
115	U:1175250.1:2000SEP08	70234155V1	711	958
115	U:1175250.1:2000SEP08	70234159V1	847	1012
115	U:1175250.1:2000SEP08	70214667V1	852	1220
115	U:1175250.1:2000SEP08	70234174V1	996	1220
115	U:1175250.1:2000SEP08	g7280771	1058	1247
115	U:1175250.1:2000SEP08	70412113D1	86	636
115	U:1175250.1:2000SEP08	70413868D1	49	640
115	U:1175250.1:2000SEP08	70410564D1	638	1168
115	U:1175250.1:2000SEP08	70391109D1	1	404
115	U:1175250.1:2000SEP08	70411412D1	407	968
115	U:1175250.1:2000SEP08	70413500D1	682	1238
115	U:1175250.1:2000SEP08	70411631D1	1	490
115	U:1175250.1:2000SEP08	70411813D1	406	865
115	U:1175250.1:2000SEP08	70413425D1	638	1080
116	U:1183192.1:2000SEP08	71088812V1	1	155
116	U:1183192.1:2000SEP08	71257773V1	746	1364
116	U:1183192.1:2000SEP08	71096474V1	750	1401
116	U:1183192.1:2000SEP08	71096558V1	953	1563
116	U:1183192.1:2000SEP08	71095372V1	415	1010
116	U:1183192.1:2000SEP08	71093830V1	555	1177
116	U:1183192.1:2000SEP08	71094309V1	616	1183

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
116	LI:1183192.1:2000SEP08	71257202V1	668	1275
116	LI:1183192.1:2000SEP08	71094327V1	565	1123
116	LI:1183192.1:2000SEP08	71093484V1	858	1513
116	LI:1183192.1:2000SEP08	71093490V1	1004	1673
116	LI:1183192.1:2000SEP08	6844754T8	56	577
116	LI:1183192.1:2000SEP08	6844754H1	60	607
116	LI:1183192.1:2000SEP08	6844754F8	61	553
116	LI:1183192.1:2000SEP08	71256250V1	134	347
116	LI:1183192.1:2000SEP08	71095732V1	302	938
116	LI:1183192.1:2000SEP08	71258218V1	583	1133
116	LI:1183192.1:2000SEP08	71095821V1	586	1228
116	LI:1183192.1:2000SEP08	71257875V1	659	1261
116	LI:1183192.1:2000SEP08	71092947V1	1	508
116	LI:1183192.1:2000SEP08	71095852V1	674	1284
116	LI:1183192.1:2000SEP08	71094229V1	684	1272
116	LI:1183192.1:2000SEP08	71093914V1	798	1447
116	LI:1183192.1:2000SEP08	7000002H1	817	1327
116	LI:1183192.1:2000SEP08	71095983V1	846	1496
116	LI:1183192.1:2000SEP08	71256947V1	905	1464
116	LI:1183192.1:2000SEP08	71093319V1	1	504
116	LI:1183192.1:2000SEP08	71094826V1	905	1422
116	LI:1183192.1:2000SEP08	71257909V1	1	553
116	LI:1183192.1:2000SEP08	4873883F7	1	551
116	LI:1183192.1:2000SEP08	70898470V1	1069	1780
116	LI:1183192.1:2000SEP08	71258120V1	1147	1739
116	LI:1183192.1:2000SEP08	4873883H1	2	264
116	LI:1183192.1:2000SEP08	71092930V1	1280	1784
116	LI:1183192.1:2000SEP08	71093620V1	1280	1922
116	LI:1183192.1:2000SEP08	2846878H1	1	262
116	LI:1183192.1:2000SEP08	2638526T6	1385	1815
116	LI:1183192.1:2000SEP08	3046268F6	1464	1815
116	LI:1183192.1:2000SEP08	3046268H1	1464	1742
116	LI:1183192.1:2000SEP08	702550H1	1494	1712
116	LI:1183192.1:2000SEP08	3046268T6	1604	1830
116	LI:1183192.1:2000SEP08	398973T6	1620	1815
116	LI:1183192.1:2000SEP08	868360H1	1624	1854
116	LI:1183192.1:2000SEP08	398973F1	1636	2148
116	LI:1183192.1:2000SEP08	g6073621	1695	1815
116	LI:1183192.1:2000SEP08	71094024V1	802	1214
116	LI:1183192.1:2000SEP08	71257395V1	878	1513
116	LI:1183192.1:2000SEP08	71094896V1	799	1195
116	LI:1183192.1:2000SEP08	71095472V1	631	1092
116	LI:1183192.1:2000SEP08	71094796V1	555	1190
116	LI:1183192.1:2000SEP08	71093537V1	329	966
116	LI:1183192.1:2000SEP08	8126105H1	1	539
116	LI:1183192.1:2000SEP08	3465420F6	1	230
116	LI:1183192.1:2000SEP08	3465420H1	1	283
116	LI:1183192.1:2000SEP08	71258561V1	1083	1724
116	LI:1183192.1:2000SEP08	71094039V1	1046	1588

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
116	U:1183192.1:2000SEP08	71257745V1	939	1565
116	U:1183192.1:2000SEP08	71256818V1	811	1399
117	U:1183325.1:2000SEP08	71076888V1	1519	2001
117	U:1183325.1:2000SEP08	71080372V1	1097	1564
117	U:1183325.1:2000SEP08	71080021V1	1698	2112
117	U:1183325.1:2000SEP08	71078890V1	864	1391
117	U:1183325.1:2000SEP08	71080042V1	987	1532
117	U:1183325.1:2000SEP08	70875127V1	1923	2113
117	U:1183325.1:2000SEP08	71560404V1	776	1421
117	U:1183325.1:2000SEP08	71559712V1	16	616
117	U:1183325.1:2000SEP08	71079939V1	794	1226
117	U:1183325.1:2000SEP08	70875618V1	936	1310
117	U:1183325.1:2000SEP08	71559647V1	964	1563
117	U:1183325.1:2000SEP08	4649460H1	1994	2274
117	U:1183325.1:2000SEP08	4592840H1	2000	2116
117	U:1183325.1:2000SEP08	4649358H1	1994	2263
117	U:1183325.1:2000SEP08	71559616V1	1	535
117	U:1183325.1:2000SEP08	6779183J1	1	341
117	U:1183325.1:2000SEP08	6779183R8	1	403
117	U:1183325.1:2000SEP08	6482105H1	1	365
117	U:1183325.1:2000SEP08	3394349H1	1	92
117	U:1183325.1:2000SEP08	3394349F6	1	391
117	U:1183325.1:2000SEP08	4407036F8	1	402
117	U:1183325.1:2000SEP08	71475254V1	15	639
117	U:1183325.1:2000SEP08	4407036H1	16	105
117	U:1183325.1:2000SEP08	5183573H1	168	305
117	U:1183325.1:2000SEP08	5183573F8	168	657
117	U:1183325.1:2000SEP08	4304365H1	478	595
117	U:1183325.1:2000SEP08	8020215J1	514	1098
117	U:1183325.1:2000SEP08	3202244F6	541	1070
117	U:1183325.1:2000SEP08	3202244H1	542	802
117	U:1183325.1:2000SEP08	71556306V1	561	1087
117	U:1183325.1:2000SEP08	5585887H1	591	665
117	U:1183325.1:2000SEP08	5585808H1	590	691
117	U:1183325.1:2000SEP08	5585887F6	590	1042
117	U:1183325.1:2000SEP08	71556462V1	624	1121
117	U:1183325.1:2000SEP08	70612109V1	642	1157
117	U:1183325.1:2000SEP08	2232645H1	642	853
117	U:1183325.1:2000SEP08	2232645F6	642	1073
117	U:1183325.1:2000SEP08	2543464H1	704	945
117	U:1183325.1:2000SEP08	71556211V1	720	1319
117	U:1183325.1:2000SEP08	71076543V1	795	1291
117	U:1183325.1:2000SEP08	60201103V1	802	1284
117	U:1183325.1:2000SEP08	71559461V1	914	1617
117	U:1183325.1:2000SEP08	70610645V1	924	1578
117	U:1183325.1:2000SEP08	71556156V1	964	1572
117	U:1183325.1:2000SEP08	8039329J1	974	1475
117	U:1183325.1:2000SEP08	71076870V1	1032	1617
117	U:1183325.1:2000SEP08	70873469V1	1123	1327



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
117	LI:1183325.1:2000SEP08	71080137V1	1443	1857
117	LI:1183325.1:2000SEP08	71078193V1	1625	2227
117	LI:1183325.1:2000SEP08	g4328477	1792	2112
117	LI:1183325.1:2000SEP08	71075117V1	1992	2264
117	LI:1183325.1:2000SEP08	71074916V1	1515	1760
117	LI:1183325.1:2000SEP08	71559508V1	1	533
117	LI:1183325.1:2000SEP08	71076271V1	1165	1683
118	LI:1178269.2:2000SEP08	60203895U1	1	473
119	LI:813422.1:2000SEP08	2497517H1	1	321
119	LI:813422.1:2000SEP08	2497517F6	1	497
119	LI:813422.1:2000SEP08	70167792V1	6	428
119	LI:813422.1:2000SEP08	70168921V1	43	556
119	LI:813422.1:2000SEP08	70165351V1	315	790
119	LI:813422.1:2000SEP08	70165894V1	334	834
119	LI:813422.1:2000SEP08	70166959V1	399	911
119	LI:813422.1:2000SEP08	70166310V1	413	885
119	LI:813422.1:2000SEP08	70169334V1	505	981
119	LI:813422.1:2000SEP08	70168807V1	636	1099
119	LI:813422.1:2000SEP08	70165598V1	727	1203
119	LI:813422.1:2000SEP08	70166092V1	772	1271
119	LI:813422.1:2000SEP08	1006417H1	868	1050
119	LI:813422.1:2000SEP08	70165798V1	899	1413
119	LI:813422.1:2000SEP08	70164520V1	960	1393
119	LI:813422.1:2000SEP08	70166061V1	1016	1543
119	LI:813422.1:2000SEP08	6526780H1	1018	1605
119	LI:813422.1:2000SEP08	70168451V1	1037	1521
119	LI:813422.1:2000SEP08	5263856H2	1049	1144
119	LI:813422.1:2000SEP08	3338768H1	1144	1385
119	LI:813422.1:2000SEP08	70166692V1	1177	1657
119	LI:813422.1:2000SEP08	466542H1	1222	1370
119	LI:813422.1:2000SEP08	6433091H1	1231	1626
119	LI:813422.1:2000SEP08	6433091T8	1313	1626
119	LI:813422.1:2000SEP08	70164598V1	1330	1781
119	LI:813422.1:2000SEP08	g1485371	1662	1802
120	LI:1093049.6:2000SEP08	5399109T9	983	1243
120	LI:1093049.6:2000SEP08	2368366H2	1016	1233
120	LI:1093049.6:2000SEP08	g2106702	1054	1410
120	LI:1093049.6:2000SEP08	g5446135	1076	1407
120	LI:1093049.6:2000SEP08	g5812267	1085	1411
120	LI:1093049.6:2000SEP08	3112876T6	1124	1364
120	LI:1093049.6:2000SEP08	2780828T6	1173	1365
120	LI:1093049.6:2000SEP08	2779312F6	1180	1407
120	LI:1093049.6:2000SEP08	2779312H1	1180	1407
120	LI:1093049.6:2000SEP08	1684340H1	1188	1407
120	LI:1093049.6:2000SEP08	810109H1	1318	1407
120	LI:1093049.6:2000SEP08	6709759H1	1	541
120	LI:1093049.6:2000SEP08	6399778F6	369	929
120	LI:1093049.6:2000SEP08	6399778H1	369	632
120	LI:1093049.6:2000SEP08	6178145H1	390	693

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
120	LI:1093049.6:2000SEP08	6178145F8	390	1037
120	LI:1093049.6:2000SEP08	g2142084	683	1062
120	LI:1093049.6:2000SEP08	6178145T8	704	1305
120	LI:1093049.6:2000SEP08	1781630H1	757	976
120	LI:1093049.6:2000SEP08	5697086H1	877	1093
120	LI:1093049.6:2000SEP08	936773T6	891	1365
120	LI:1093049.6:2000SEP08	629645H1	916	1023
120	LI:1093049.6:2000SEP08	2611430H1	917	1188
120	LI:1093049.6:2000SEP08	2651724H1	917	1119
120	LI:1093049.6:2000SEP08	g2767207	941	1414
120	LI:1093049.6:2000SEP08	g3898169	944	1402
120	LI:1093049.6:2000SEP08	g4393504	950	1408
120	LI:1093049.6:2000SEP08	g5393320	971	1407
121	LI:202192.4:2000SEP08	6479780F9	1	515
121	LI:202192.4:2000SEP08	6479780H1	1	350
121	LI:202192.4:2000SEP08	6777378J1	1	649
122	LG:1041854.1:2000SEP08	6798561H1	168	688
122	LG:1041854.1:2000SEP08	6797277H1	168	747
122	LG:1041854.1:2000SEP08	6793190T8	254	802
122	LG:1041854.1:2000SEP08	6798561T8	261	783
122	LG:1041854.1:2000SEP08	6797887T8	274	777
122	LG:1041854.1:2000SEP08	6797887H1	1	540
122	LG:1041854.1:2000SEP08	6797887F8	1	511
122	LG:1041854.1:2000SEP08	6793190F8	168	756
122	LG:1041854.1:2000SEP08	6797277F8	168	770
122	LG:1041854.1:2000SEP08	6798561F8	168	837
122	LG:1041854.1:2000SEP08	6793190H1	168	691
122	LG:1041854.1:2000SEP08	6797277T8	319	754
122	LG:1041854.1:2000SEP08	g6700685	416	882
123	LG:1100502.1:2000SEP08	6793312H1	2	451
123	LG:1100502.1:2000SEP08	6797957H1	1	449
123	LG:1100502.1:2000SEP08	6793583H1	46	504
123	LG:1100502.1:2000SEP08	6798145H1	46	477
123	LG:1100502.1:2000SEP08	g1685926	121	439
124	LI:726414.1:2000SEP08	6268429H1	1	444
124	LI:726414.1:2000SEP08	6268429F8	119	668
124	LI:726414.1:2000SEP08	6268429T8	265	925
125	LI:400517.4:2000SEP08	7336884H1	3108	3523
125	LI:400517.4:2000SEP08	71345951V1	2452	3035
125	LI:400517.4:2000SEP08	2669866F6	2658	2992
125	LI:400517.4:2000SEP08	7674612J1	2345	2529
125	LI:400517.4:2000SEP08	70900352V1	2352	2851
125	LI:400517.4:2000SEP08	71269692V1	2351	3000
125	LI:400517.4:2000SEP08	71090658V1	2368	2941
125	LI:400517.4:2000SEP08	70900660V1	2402	2955
125	LI:400517.4:2000SEP08	70900544V1	2404	2939
125	LI:400517.4:2000SEP08	71270668V1	2428	2909
125	LI:400517.4:2000SEP08	71346046V1	2439	2964
125	LI:400517.4:2000SEP08	71269735V1	2418	2962

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
125	LI:400517.4:2000SEP08	70936051V1	2714	2988
125	LI:400517.4:2000SEP08	7440147H1	2714	2970
125	LI:400517.4:2000SEP08	70525048V1	1670	1894
125	LI:400517.4:2000SEP08	7978404H1	1	578
125	LI:400517.4:2000SEP08	7978442H1	1	595
125	LI:400517.4:2000SEP08	7440568H1	379	865
125	LI:400517.4:2000SEP08	2343349F6	1150	1536
125	LI:400517.4:2000SEP08	2343349H1	1150	1403
125	LI:400517.4:2000SEP08	70937660V1	1150	1615
125	LI:400517.4:2000SEP08	7986155H2	455	1106
125	LI:400517.4:2000SEP08	6470963H1	697	1226
125	LI:400517.4:2000SEP08	012910H1	712	977
125	LI:400517.4:2000SEP08	70646110V1	1173	1665
125	LI:400517.4:2000SEP08	6488039H1	1194	1725
125	LI:400517.4:2000SEP08	3739117H1	1208	1371
125	LI:400517.4:2000SEP08	70523411V1	1261	1878
125	LI:400517.4:2000SEP08	8033302H1	1290	1924
125	LI:400517.4:2000SEP08	6470963F8	711	1257
125	LI:400517.4:2000SEP08	70523527V1	1290	1769
125	LI:400517.4:2000SEP08	g1775861	788	1176
125	LI:400517.4:2000SEP08	258893R1	846	1333
125	LI:400517.4:2000SEP08	70523309V1	858	1550
125	LI:400517.4:2000SEP08	70524645V1	858	1462
125	LI:400517.4:2000SEP08	2633233F6	858	1263
125	LI:400517.4:2000SEP08	2633233H1	858	1142
125	LI:400517.4:2000SEP08	70535526V1	896	1152
125	LI:400517.4:2000SEP08	7608962H1	1010	1628
125	LI:400517.4:2000SEP08	7608962J1	1036	1616
125	LI:400517.4:2000SEP08	6354826H1	1141	1436
125	LI:400517.4:2000SEP08	70524958V1	1153	1727
125	LI:400517.4:2000SEP08	3489591H1	1402	1692
125	LI:400517.4:2000SEP08	70526475V1	1556	2291
125	LI:400517.4:2000SEP08	258893F1	2714	2996
125	LI:400517.4:2000SEP08	g1775807	2714	3003
125	LI:400517.4:2000SEP08	71087946V1	2350	2815
125	LI:400517.4:2000SEP08	4597511H1	2231	2321
125	LI:400517.4:2000SEP08	71090729V1	2327	2956
125	LI:400517.4:2000SEP08	71090386V1	2333	2659
125	LI:400517.4:2000SEP08	70524730V1	1601	2306
125	LI:400517.4:2000SEP08	7439352H1	1659	2276
125	LI:400517.4:2000SEP08	71270813V1	2329	2915
125	LI:400517.4:2000SEP08	70911734V1	2350	2913
125	LI:400517.4:2000SEP08	2460346H1	2692	2850
125	LI:400517.4:2000SEP08	2669866H1	2658	2940
125	LI:400517.4:2000SEP08	2669866T6	2655	2952
125	LI:400517.4:2000SEP08	g993716	2680	2991
125	LI:400517.4:2000SEP08	6915443H1	2693	3190
125	LI:400517.4:2000SEP08	1701274F6	2683	3194
125	LI:400517.4:2000SEP08	1701274H1	2683	2940

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
125	LI:400517.4:2000SEP08	1384769H1	2692	2833
125	LI:400517.4:2000SEP08	g6036967	2714	3003
125	LI:400517.4:2000SEP08	2343349T6	2692	2958
125	LI:400517.4:2000SEP08	5692896H1	2692	2931
125	LI:400517.4:2000SEP08	g1939622	2710	2996
125	LI:400517.4:2000SEP08	70938932V1	2714	3004
125	LI:400517.4:2000SEP08	g4109540	2714	3003
125	LI:400517.4:2000SEP08	g4510275	2714	2997
125	LI:400517.4:2000SEP08	70937828V1	2714	2988
125	LI:400517.4:2000SEP08	2382147H1	2978	3194
125	LI:400517.4:2000SEP08	70936534V1	2230	2912
125	LI:400517.4:2000SEP08	2633233T6	2714	2958
125	LI:400517.4:2000SEP08	70523444V1	2714	2988
125	LI:400517.4:2000SEP08	3926162H1	2714	2948
125	LI:400517.4:2000SEP08	70935030V1	2122	2490
125	LI:400517.4:2000SEP08	70936441V1	2104	2726
125	LI:400517.4:2000SEP08	g6044900	2941	2996
125	LI:400517.4:2000SEP08	70935467V1	1876	2321
125	LI:400517.4:2000SEP08	70646141V1	1948	2214
125	LI:400517.4:2000SEP08	g2063236	1871	2320
125	LI:400517.4:2000SEP08	70936928V1	1957	2321
125	LI:400517.4:2000SEP08	70913530V1	2335	3031
125	LI:400517.4:2000SEP08	2685985H1	2793	3001
125	LI:400517.4:2000SEP08	g2432619	2855	2982
126	LI:1078917.1:2000SEP08	6790926F8	1	661
126	LI:1078917.1:2000SEP08	6790926H1	1	173
126	LI:1078917.1:2000SEP08	6790926T8	330	978
126	LI:1078917.1:2000SEP08	6792520H1	568	1053
126	LI:1078917.1:2000SEP08	6792520F8	568	1084
126	LI:1078917.1:2000SEP08	6792520T8	568	979
126	LI:1078917.1:2000SEP08	6797488H1	570	1046
126	LI:1078917.1:2000SEP08	6797488F8	713	1110
127	LI:1012560.1:2000SEP08	71420033V1	600	661
127	LI:1012560.1:2000SEP08	6982025H1	620	1147
127	LI:1012560.1:2000SEP08	71433170V1	1	656
127	LI:1012560.1:2000SEP08	71433478V1	5	643
127	LI:1012560.1:2000SEP08	275757R6	36	402
127	LI:1012560.1:2000SEP08	275757H1	36	298
127	LI:1012560.1:2000SEP08	71433178V1	36	727
127	LI:1012560.1:2000SEP08	275757T6	137	629
127	LI:1012560.1:2000SEP08	g4265765	200	599
127	LI:1012560.1:2000SEP08	3220669H1	316	604
127	LI:1012560.1:2000SEP08	71422565V1	325	541
127	LI:1012560.1:2000SEP08	6982025R8	331	905
127	LI:1012560.1:2000SEP08	71440747V1	346	550
127	LI:1012560.1:2000SEP08	71422738V1	349	521
127	LI:1012560.1:2000SEP08	71466675V1	361	530
127	LI:1012560.1:2000SEP08	70253156V1	442	517
127	LI:1012560.1:2000SEP08	6982025F8	519	1147

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
128	LI:427997.4:2000SEP08	g2969311	1954	2238
128	LI:427997.4:2000SEP08	1853191F6	1722	2234
128	LI:427997.4:2000SEP08	1853191H1	1722	1976
128	LI:427997.4:2000SEP08	3964084H1	1018	1310
128	LI:427997.4:2000SEP08	70059533V1	1338	1601
128	LI:427997.4:2000SEP08	g1981676	1338	1670
128	LI:427997.4:2000SEP08	7059995H1	802	1287
128	LI:427997.4:2000SEP08	5376913H1	683	938
128	LI:427997.4:2000SEP08	1683557F6	1687	2243
128	LI:427997.4:2000SEP08	70061871V1	1816	2167
128	LI:427997.4:2000SEP08	g5858285	1917	2240
128	LI:427997.4:2000SEP08	3409241T6	1927	2383
128	LI:427997.4:2000SEP08	5669576H1	1947	2137
128	LI:427997.4:2000SEP08	3628624H1	1952	2267
128	LI:427997.4:2000SEP08	g5855887	1917	2239
128	LI:427997.4:2000SEP08	2408847H1	1898	2148
128	LI:427997.4:2000SEP08	70059242V1	1904	2238
128	LI:427997.4:2000SEP08	960516H1	1875	2176
128	LI:427997.4:2000SEP08	960387T1	1875	2201
128	LI:427997.4:2000SEP08	2086141H1	1880	2143
128	LI:427997.4:2000SEP08	70062540V1	1876	2238
128	LI:427997.4:2000SEP08	3751266T6	1741	2203
128	LI:427997.4:2000SEP08	3321818T6	1746	2201
128	LI:427997.4:2000SEP08	2298374T6	1751	2198
128	LI:427997.4:2000SEP08	70527491V1	854	1515
128	LI:427997.4:2000SEP08	71265279V1	898	1591
128	LI:427997.4:2000SEP08	4239442T8	1900	2310
128	LI:427997.4:2000SEP08	71265775V1	804	1384
128	LI:427997.4:2000SEP08	4177726H1	811	1115
128	LI:427997.4:2000SEP08	g2541182	2122	2440
128	LI:427997.4:2000SEP08	g3039868	2163	2241
128	LI:427997.4:2000SEP08	223027H1	2079	2238
128	LI:427997.4:2000SEP08	223027F1	2080	2238
128	LI:427997.4:2000SEP08	70524882V1	1575	1741
128	LI:427997.4:2000SEP08	71120713V1	1322	1901
128	LI:427997.4:2000SEP08	71118117V1	1327	1887
128	LI:427997.4:2000SEP08	71118747V1	1311	1976
128	LI:427997.4:2000SEP08	70527304V1	1317	1985
128	LI:427997.4:2000SEP08	2298374H1	1639	1911
128	LI:427997.4:2000SEP08	1361272F1	1673	2232
128	LI:427997.4:2000SEP08	71119759V1	731	1250
128	LI:427997.4:2000SEP08	71266124V1	732	1382
128	LI:427997.4:2000SEP08	71118641V1	731	1327
128	LI:427997.4:2000SEP08	70530256V1	467	1053
128	LI:427997.4:2000SEP08	1915736R6	731	1151
128	LI:427997.4:2000SEP08	71266413V1	932	1624
128	LI:427997.4:2000SEP08	g1980460	2011	2329
128	LI:427997.4:2000SEP08	223027R1	2080	2238
128	LI:427997.4:2000SEP08	3934758F6	904	1427

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
128	LI:427997.4:2000SEP08	096309H1	1114	1339
128	LI:427997.4:2000SEP08	71118507V1	1131	1716
128	LI:427997.4:2000SEP08	71120385V1	1134	1464
128	LI:427997.4:2000SEP08	5947885H1	1113	1379
128	LI:427997.4:2000SEP08	71266010V1	1110	1781
128	LI:427997.4:2000SEP08	4717156H1	1293	1517
128	LI:427997.4:2000SEP08	5391806H1	1298	1497
128	LI:427997.4:2000SEP08	1623591F6	1869	2238
128	LI:427997.4:2000SEP08	1623591H1	1869	2090
128	LI:427997.4:2000SEP08	g2198272	1874	2236
128	LI:427997.4:2000SEP08	960516R1	1875	2238
128	LI:427997.4:2000SEP08	70061519V1	1763	2349
128	LI:427997.4:2000SEP08	3021149H1	1774	2063
128	LI:427997.4:2000SEP08	g2879046	1791	2241
128	LI:427997.4:2000SEP08	g2444573	1810	2237
128	LI:427997.4:2000SEP08	70062956V1	1809	2106
128	LI:427997.4:2000SEP08	g6656753	1811	2238
128	LI:427997.4:2000SEP08	71265047V1	731	1277
128	LI:427997.4:2000SEP08	1853191T6	1797	2392
128	LI:427997.4:2000SEP08	70061837V1	1851	2238
128	LI:427997.4:2000SEP08	g6975043	1856	2239
128	LI:427997.4:2000SEP08	71266623V1	1464	1767
128	LI:427997.4:2000SEP08	71265365V1	1463	1908
128	LI:427997.4:2000SEP08	1361366H1	1673	1850
128	LI:427997.4:2000SEP08	71118731V1	731	1308
128	LI:427997.4:2000SEP08	71265792V1	731	1276
128	LI:427997.4:2000SEP08	71119741V1	984	1468
128	LI:427997.4:2000SEP08	5287845H1	1333	1456
128	LI:427997.4:2000SEP08	71264971V1	722	1497
128	LI:427997.4:2000SEP08	g2198304	1147	1236
128	LI:427997.4:2000SEP08	3931985H1	905	1196
128	LI:427997.4:2000SEP08	71120838V1	1291	1757
128	LI:427997.4:2000SEP08	2681128H1	696	995
128	LI:427997.4:2000SEP08	70531848V1	706	1276
128	LI:427997.4:2000SEP08	71266663V1	731	1355
128	LI:427997.4:2000SEP08	2681128F6	695	1194
128	LI:427997.4:2000SEP08	g850466	493	827
128	LI:427997.4:2000SEP08	2560627H1	650	920
128	LI:427997.4:2000SEP08	2557738H1	650	901
128	LI:427997.4:2000SEP08	70528634V1	558	1219
128	LI:427997.4:2000SEP08	1683573T6	1741	2191
128	LI:427997.4:2000SEP08	70060889V1	1831	2284
128	LI:427997.4:2000SEP08	1623591T6	1835	2193
128	LI:427997.4:2000SEP08	1915736T6	1834	2192
128	LI:427997.4:2000SEP08	g3155476	1837	2224
128	LI:427997.4:2000SEP08	g5392637	1842	2238
128	LI:427997.4:2000SEP08	4778510H1	1818	2096
128	LI:427997.4:2000SEP08	70061698V1	1827	2238
128	LI:427997.4:2000SEP08	g4901915	1832	2238

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
128	LI:427997.4:2000SEP08	71118652V1	1424	2039
128	LI:427997.4:2000SEP08	3934357H1	904	1216
128	LI:427997.4:2000SEP08	1336611H1	1714	1957
128	LI:427997.4:2000SEP08	5290693H1	1330	1626
128	LI:427997.4:2000SEP08	70060883V1	1338	1745
128	LI:427997.4:2000SEP08	71117317V1	1166	1740
128	LI:427997.4:2000SEP08	71119085V1	1213	1841
128	LI:427997.4:2000SEP08	3792653H1	1240	1449
128	LI:427997.4:2000SEP08	6329373H1	1251	1796
128	LI:427997.4:2000SEP08	70062844V1	1485	1881
128	LI:427997.4:2000SEP08	2273538H1	1510	1779
128	LI:427997.4:2000SEP08	3022474H1	1556	1849
128	LI:427997.4:2000SEP08	71120487V1	1115	1610
128	LI:427997.4:2000SEP08	71265619V1	1043	1705
128	LI:427997.4:2000SEP08	5946829H1	1056	1341
128	LI:427997.4:2000SEP08	71266345V1	1110	1765
128	LI:427997.4:2000SEP08	71120506V1	1420	1873
128	LI:427997.4:2000SEP08	71119165V1	893	1508
128	LI:427997.4:2000SEP08	3173959T6	1600	2191
128	LI:427997.4:2000SEP08	1915736H1	731	980
128	LI:427997.4:2000SEP08	71265549V1	731	1218
128	LI:427997.4:2000SEP08	g3151378	1976	2242
128	LI:427997.4:2000SEP08	71119545V1	1283	1970
128	LI:427997.4:2000SEP08	3173959H1	1338	1605
128	LI:427997.4:2000SEP08	1683573F6	1687	2102
128	LI:427997.4:2000SEP08	1683573H1	1687	1923
128	LI:427997.4:2000SEP08	1632340H1	1721	1938
128	LI:427997.4:2000SEP08	614199H1	1722	1949
128	LI:427997.4:2000SEP08	70059326V1	1735	2238
128	LI:427997.4:2000SEP08	2298374R6	1639	2093
128	LI:427997.4:2000SEP08	3408296H1	1578	1828
128	LI:427997.4:2000SEP08	71119740V1	1148	1440
128	LI:427997.4:2000SEP08	4174878H1	1152	1436
128	LI:427997.4:2000SEP08	3751266F6	1	358
128	LI:427997.4:2000SEP08	71266530V1	732	1262
128	LI:427997.4:2000SEP08	71264808V1	753	1279
128	LI:427997.4:2000SEP08	5296714H1	765	1049
128	LI:427997.4:2000SEP08	71120676V1	1562	1948
128	LI:427997.4:2000SEP08	70529734V1	461	579
128	LI:427997.4:2000SEP08	3751266H1	1	296
128	LI:427997.4:2000SEP08	3321818F6	103	497
128	LI:427997.4:2000SEP08	3321818H1	104	350
128	LI:427997.4:2000SEP08	6351045H2	329	660
128	LI:427997.4:2000SEP08	3662984H1	1018	1270
128	LI:427997.4:2000SEP08	70529227V1	1021	1617
128	LI:427997.4:2000SEP08	70061612V1	1338	1898
128	LI:427997.4:2000SEP08	70059581V1	1338	1831
128	LI:427997.4:2000SEP08	70058920V1	1338	1778
128	LI:427997.4:2000SEP08	70530468V1	1336	1867

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
128	LI:427997.4:2000SEP08	3173959F6	1338	1902
128	LI:427997.4:2000SEP08	1865168H1	1364	1621
128	LI:427997.4:2000SEP08	71294927V1	1419	1832
128	LI:427997.4:2000SEP08	70062034V1	1342	1810
128	LI:427997.4:2000SEP08	70062398V1	1357	1878
128	LI:427997.4:2000SEP08	71118139V1	1362	1925
128	LI:427997.4:2000SEP08	71265724V1	731	1025
128	LI:427997.4:2000SEP08	3934758H1	904	1209
128	LI:427997.4:2000SEP08	3934538H1	904	1216
128	LI:427997.4:2000SEP08	71117117V1	931	1596
128	LI:427997.4:2000SEP08	3672409H1	947	1180
128	LI:427997.4:2000SEP08	71118560V1	979	1553
129	LI:197899.1:2000SEP08	g1855077	1843	2224
129	LI:197899.1:2000SEP08	g4740828	1847	2224
129	LI:197899.1:2000SEP08	g5755916	1813	2225
129	LI:197899.1:2000SEP08	g5755042	1827	2224
129	LI:197899.1:2000SEP08	g4283562	1832	2218
129	LI:197899.1:2000SEP08	g6698570	1810	2117
129	LI:197899.1:2000SEP08	7160288H1	1	417
129	LI:197899.1:2000SEP08	60211613U1	1	533
129	LI:197899.1:2000SEP08	7251266H1	1	532
129	LI:197899.1:2000SEP08	3584679F6	4	290
129	LI:197899.1:2000SEP08	3584679H1	4	124
129	LI:197899.1:2000SEP08	3174625H1	12	248
129	LI:197899.1:2000SEP08	7251266F8	16	553
129	LI:197899.1:2000SEP08	60211616U1	84	573
129	LI:197899.1:2000SEP08	70300001D1	89	527
129	LI:197899.1:2000SEP08	7711767H2	176	708
129	LI:197899.1:2000SEP08	7749453H1	416	861
129	LI:197899.1:2000SEP08	70300067D1	469	811
129	LI:197899.1:2000SEP08	268719H1	531	854
129	LI:197899.1:2000SEP08	7711767J1	761	1207
129	LI:197899.1:2000SEP08	7961479H1	860	1403
129	LI:197899.1:2000SEP08	1273037T6	2090	2153
129	LI:197899.1:2000SEP08	5022828H1	1117	1382
129	LI:197899.1:2000SEP08	2416733H1	1151	1379
129	LI:197899.1:2000SEP08	70563496V1	1209	1438
129	LI:197899.1:2000SEP08	70564238V1	1273	1827
129	LI:197899.1:2000SEP08	70563282V1	1341	2006
129	LI:197899.1:2000SEP08	70564302V1	1357	1935
129	LI:197899.1:2000SEP08	3278713H1	1495	1746
129	LI:197899.1:2000SEP08	7251266T8	1553	2125
129	LI:197899.1:2000SEP08	70553617V1	1554	1672
129	LI:197899.1:2000SEP08	5022828T9	1557	1896
129	LI:197899.1:2000SEP08	70566741V1	1573	2056
129	LI:197899.1:2000SEP08	70563828V1	1609	1661
129	LI:197899.1:2000SEP08	70565215V1	1747	2195
129	LI:197899.1:2000SEP08	g1803535	1764	2219
129	LI:197899.1:2000SEP08	g5764951	1894	2216



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
129	LI:197899.1:2000SEP08	70300023D1	2035	2564
129	LI:197899.1:2000SEP08	5022828F8	1103	1573
129	LI:197899.1:2000SEP08	1687189H1	1055	1243
129	LI:197899.1:2000SEP08	1273037H1	1059	1297
129	LI:197899.1:2000SEP08	1273037F1	1059	1574
129	LI:197899.1:2000SEP08	1273037F6	1059	1374
130	LG:334199.1:2000SEP08	g2188226	51	490
130	LG:334199.1:2000SEP08	g1925235	185	320
130	LG:334199.1:2000SEP08	5287781H1	1	282
130	LG:334199.1:2000SEP08	1237250H1	138	267
130	LG:334199.1:2000SEP08	6825741J1	118	747
130	LG:334199.1:2000SEP08	g1933781	185	542
130	LG:334199.1:2000SEP08	g1856226	59	507
130	LG:334199.1:2000SEP08	g6698815	833	1003
130	LG:334199.1:2000SEP08	6307741H1	382	882
130	LG:334199.1:2000SEP08	g2594186	473	784
130	LG:334199.1:2000SEP08	g3281524	627	1104
130	LG:334199.1:2000SEP08	6766090J1	491	1095
130	LG:334199.1:2000SEP08	g3331386	620	1075
130	LG:334199.1:2000SEP08	g4287817	613	1071
130	LG:334199.1:2000SEP08	g3255280	636	1063
130	LG:334199.1:2000SEP08	g1856142	707	1063
130	LG:334199.1:2000SEP08	g1925119	737	1063
130	LG:334199.1:2000SEP08	g1933726	694	1063
130	LG:334199.1:2000SEP08	g2240977	657	1034
130	LG:334199.1:2000SEP08	g6710093	607	1018
130	LG:334199.1:2000SEP08	g3240735	1097	1478
130	LG:334199.1:2000SEP08	6825741H1	757	1218
131	LG:334345.1:2000SEP08	2108692T6	1	606
131	LG:334345.1:2000SEP08	2108692R6	7	463
131	LG:334345.1:2000SEP08	2108692H1	7	240
132	LG:228092.1:2000SEP08	g4629927	960	1289
132	LG:228092.1:2000SEP08	g761683	1024	1279
132	LG:228092.1:2000SEP08	4360345H1	1038	1286
132	LG:228092.1:2000SEP08	g777314	1047	1276
132	LG:228092.1:2000SEP08	g2957391	947	1281
132	LG:228092.1:2000SEP08	g6439036	957	1281
132	LG:228092.1:2000SEP08	g2444549	860	1277
132	LG:228092.1:2000SEP08	2005241R6	747	1275
132	LG:228092.1:2000SEP08	1974473H1	781	1043
132	LG:228092.1:2000SEP08	g777313	792	1018
132	LG:228092.1:2000SEP08	348495T6	812	1255
132	LG:228092.1:2000SEP08	g5630955	825	1277
132	LG:228092.1:2000SEP08	5807224H1	827	1063
132	LG:228092.1:2000SEP08	5886974H1	827	1087
132	LG:228092.1:2000SEP08	g2706302	827	1277
132	LG:228092.1:2000SEP08	5807356H1	827	1047
132	LG:228092.1:2000SEP08	5881415H1	828	1111
132	LG:228092.1:2000SEP08	g5904087	835	1277

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
132	LG:228092.1:2000SEP08	g4686709	845	1286
132	LG:228092.1:2000SEP08	g5394855	845	1281
132	LG:228092.1:2000SEP08	4447566H1	1	271
132	LG:228092.1:2000SEP08	3127960H1	52	323
132	LG:228092.1:2000SEP08	569430H1	54	326
132	LG:228092.1:2000SEP08	4343591H1	55	334
132	LG:228092.1:2000SEP08	5946123H1	64	295
132	LG:228092.1:2000SEP08	5947230H1	65	372
132	LG:228092.1:2000SEP08	5995210F9	79	492
132	LG:228092.1:2000SEP08	267547H1	135	307
132	LG:228092.1:2000SEP08	5899849F8	154	677
132	LG:228092.1:2000SEP08	5899849H1	154	442
132	LG:228092.1:2000SEP08	7001476H1	165	676
132	LG:228092.1:2000SEP08	5493806H1	446	708
132	LG:228092.1:2000SEP08	2326719H1	603	827
132	LG:228092.1:2000SEP08	7623093H1	656	1200
132	LG:228092.1:2000SEP08	7451259T2	672	1181
132	LG:228092.1:2000SEP08	5899849T8	680	1116
132	LG:228092.1:2000SEP08	2005241H1	747	988
132	LG:228092.1:2000SEP08	2005241T6	747	1243
133	LG:098580.1:2000SEP08	2824311F6	334	801
133	LG:098580.1:2000SEP08	2824311H1	314	609
133	LG:098580.1:2000SEP08	g3933408	94	574
133	LG:098580.1:2000SEP08	g2728655	167	574
133	LG:098580.1:2000SEP08	g3933291	476	574
133	LG:098580.1:2000SEP08	2458113T6	1	536
134	LG:969572.1:2000SEP08	5334677F8	13	600
134	LG:969572.1:2000SEP08	5334677T8	198	709
134	LG:969572.1:2000SEP08	5334677H1	1	143
135	LG:196958.1:2000SEP08	405856T6	1280	1717
135	LG:196958.1:2000SEP08	1839940H1	1279	1523
135	LG:196958.1:2000SEP08	405856H1	1280	1501
135	LG:196958.1:2000SEP08	286780F1	1270	1755
135	LG:196958.1:2000SEP08	1391412T6	1229	1641
135	LG:196958.1:2000SEP08	322252H1	1264	1507
135	LG:196958.1:2000SEP08	286780H1	1275	1630
135	LG:196958.1:2000SEP08	319439R6	907	1427
135	LG:196958.1:2000SEP08	319439H1	907	1274
135	LG:196958.1:2000SEP08	288768H1	907	1245
135	LG:196958.1:2000SEP08	2436043H1	952	1160
135	LG:196958.1:2000SEP08	2436279H1	1353	1589
135	LG:196958.1:2000SEP08	1837917H1	1143	1410
135	LG:196958.1:2000SEP08	2437258H1	629	844
135	LG:196958.1:2000SEP08	1838327T6	1091	1650
135	LG:196958.1:2000SEP08	525826H1	1113	1358
135	LG:196958.1:2000SEP08	323143H1	1296	1518
135	LG:196958.1:2000SEP08	405856R6	1280	1747
135	LG:196958.1:2000SEP08	1837332H1	860	1070
135	LG:196958.1:2000SEP08	319439R1	907	1301

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
135	LG:196958.1:2000SEP08	2438255H1	1069	1299
135	LG:196958.1:2000SEP08	319439T6	1085	1718
135	LG:196958.1:2000SEP08	323876H1	1521	1756
135	LG:196958.1:2000SEP08	g2327938	1415	1545
135	LG:196958.1:2000SEP08	320704H1	1357	1730
135	LG:196958.1:2000SEP08	408512H1	1357	1594
135	LG:196958.1:2000SEP08	319439F1	1358	1755
135	LG:196958.1:2000SEP08	319684H1	1381	1756
135	LG:196958.1:2000SEP08	337238H1	1648	1755
135	LG:196958.1:2000SEP08	2440148H1	1661	1755
135	LG:196958.1:2000SEP08	408600R6	1597	1754
135	LG:196958.1:2000SEP08	408717H1	1648	1759
135	LG:196958.1:2000SEP08	320126H1	950	1293
135	LG:196958.1:2000SEP08	2437058H1	1147	1401
135	LG:196958.1:2000SEP08	1839752H1	1164	1454
135	LG:196958.1:2000SEP08	408600H1	1597	1754
135	LG:196958.1:2000SEP08	527481H1	239	494
135	LG:196958.1:2000SEP08	432792H1	246	545
135	LG:196958.1:2000SEP08	2438229H1	240	302
135	LG:196958.1:2000SEP08	525151H1	60	324
135	LG:196958.1:2000SEP08	431103H1	113	346
135	LG:196958.1:2000SEP08	524880H1	117	332
135	LG:196958.1:2000SEP08	1389845H1	1	240
135	LG:196958.1:2000SEP08	g1989365	635	899
135	LG:196958.1:2000SEP08	1839230H1	1164	1431
135	LG:196958.1:2000SEP08	1391412H1	222	470
135	LG:196958.1:2000SEP08	1389744H1	222	475
135	LG:196958.1:2000SEP08	525757H1	210	570
135	LG:196958.1:2000SEP08	2432069H1	187	413
135	LG:196958.1:2000SEP08	431845H1	162	398
135	LG:196958.1:2000SEP08	1838327F6	121	646
135	LG:196958.1:2000SEP08	1838327H1	121	389
135	LG:196958.1:2000SEP08	1391450H1	123	254
135	LG:196958.1:2000SEP08	657629H1	144	306
135	LG:196958.1:2000SEP08	2439208H1	553	799
135	LG:196958.1:2000SEP08	1840062H1	443	711
135	LG:196958.1:2000SEP08	654194H1	280	544
135	LG:196958.1:2000SEP08	1391277H1	424	664
136	LG:1087811.1:2000SEP08	565688R6	7	291
136	LG:1087811.1:2000SEP08	565688H1	7	229
136	LG:1087811.1:2000SEP08	5199726H1	7	270
136	LG:1087811.1:2000SEP08	6167719H1	13	491
136	LG:1087811.1:2000SEP08	3396040H1	13	68
136	LG:1087811.1:2000SEP08	7655877H1	13	67
136	LG:1087811.1:2000SEP08	338433H1	13	82
136	LG:1087811.1:2000SEP08	853227R6	68	574
136	LG:1087811.1:2000SEP08	518551R1	74	547
136	LG:1087811.1:2000SEP08	7740966H1	79	722
136	LG:1087811.1:2000SEP08	339831H1	105	341

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
136	LG:1087811.1:2000SEP08	500238F1	128	728
136	LG:1087811.1:2000SEP08	342750H1	141	362
136	LG:1087811.1:2000SEP08	2919416H2	147	394
136	LG:1087811.1:2000SEP08	4441395H1	165	281
136	LG:1087811.1:2000SEP08	376238H1	187	432
136	LG:1087811.1:2000SEP08	2352459H1	213	450
136	LG:1087811.1:2000SEP08	2351409H1	213	443
136	LG:1087811.1:2000SEP08	2950735H1	648	920
136	LG:1087811.1:2000SEP08	5043837H1	666	930
136	LG:1087811.1:2000SEP08	g5744437	761	1003
136	LG:1087811.1:2000SEP08	1223772H1	784	1033
136	LG:1087811.1:2000SEP08	853227T6	803	1019
136	LG:1087811.1:2000SEP08	3624479H1	815	1019
136	LG:1087811.1:2000SEP08	7658126J1	826	1338
136	LG:1087811.1:2000SEP08	2972940T6	841	1019
136	LG:1087811.1:2000SEP08	2304130H1	539	798
136	LG:1087811.1:2000SEP08	g5362427	542	998
136	LG:1087811.1:2000SEP08	6387991H1	555	718
136	LG:1087811.1:2000SEP08	2829325H1	606	880
136	LG:1087811.1:2000SEP08	6590088H1	626	718
136	LG:1087811.1:2000SEP08	6252361H1	1	249
136	LG:1087811.1:2000SEP08	g758862	1	170
136	LG:1087811.1:2000SEP08	939450H1	1	138
136	LG:1087811.1:2000SEP08	1875744H1	1	254
136	LG:1087811.1:2000SEP08	1873658H1	1	273
136	LG:1087811.1:2000SEP08	345602H1	1	222
136	LG:1087811.1:2000SEP08	3563357H1	6	256
136	LG:1087811.1:2000SEP08	565688R1	7	459
136	LG:1087811.1:2000SEP08	g2264487	952	1019
136	LG:1087811.1:2000SEP08	146431H1	883	1074
136	LG:1087811.1:2000SEP08	144709H1	889	1015
136	LG:1087811.1:2000SEP08	144709R1	890	1019
136	LG:1087811.1:2000SEP08	1942610H1	845	1015
136	LG:1087811.1:2000SEP08	1942610R6	845	1019
136	LG:1087811.1:2000SEP08	1447494H1	849	1003
136	LG:1087811.1:2000SEP08	144709F1	878	1019
136	LG:1087811.1:2000SEP08	1942610T6	845	1019
136	LG:1087811.1:2000SEP08	6252269H1	1	317
136	LG:1087811.1:2000SEP08	375720H1	27	292
136	LG:1087811.1:2000SEP08	5340183F8	15	247
136	LG:1087811.1:2000SEP08	5805876H1	28	235
136	LG:1087811.1:2000SEP08	518551H1	50	190
136	LG:1087811.1:2000SEP08	3489279H1	51	315
136	LG:1087811.1:2000SEP08	5340183T8	58	630
136	LG:1087811.1:2000SEP08	6252361T8	19	629
136	LG:1087811.1:2000SEP08	4848110H1	65	154
136	LG:1087811.1:2000SEP08	4803270H1	21	290
136	LG:1087811.1:2000SEP08	500238H1	22	227
136	LG:1087811.1:2000SEP08	564726H1	22	241

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
136	LG:1087811.1:2000SEP08	500238R1	22	421
136	LG:1087811.1:2000SEP08	853227H1	68	315
136	LG:1087811.1:2000SEP08	518551F1	241	728
136	LG:1087811.1:2000SEP08	1746545H1	243	511
136	LG:1087811.1:2000SEP08	4878173H1	243	507
136	LG:1087811.1:2000SEP08	497017H1	256	448
136	LG:1087811.1:2000SEP08	3121516H1	257	542
136	LG:1087811.1:2000SEP08	g5837313	275	727
136	LG:1087811.1:2000SEP08	g2704490	278	728
136	LG:1087811.1:2000SEP08	6725179F8	298	852
136	LG:1087811.1:2000SEP08	g1099349	299	728
136	LG:1087811.1:2000SEP08	498297H1	299	436
136	LG:1087811.1:2000SEP08	g2279564	301	730
136	LG:1087811.1:2000SEP08	g6398356	301	728
136	LG:1087811.1:2000SEP08	6725179H1	298	852
136	LG:1087811.1:2000SEP08	g5395790	305	731
136	LG:1087811.1:2000SEP08	565688F1	214	728
136	LG:1087811.1:2000SEP08	2766321H1	217	459
136	LG:1087811.1:2000SEP08	501253H1	236	396
136	LG:1087811.1:2000SEP08	6484730H1	236	644
136	LG:1087811.1:2000SEP08	g6657428	326	718
136	LG:1087811.1:2000SEP08	341786H1	331	530
136	LG:1087811.1:2000SEP08	2972940H1	360	652
136	LG:1087811.1:2000SEP08	478130H1	370	643
136	LG:1087811.1:2000SEP08	565688T6	364	684
136	LG:1087811.1:2000SEP08	1221285H1	392	623
136	LG:1087811.1:2000SEP08	g2953838	407	731
136	LG:1087811.1:2000SEP08	563428H1	421	650
136	LG:1087811.1:2000SEP08	7658126H1	420	901
136	LG:1087811.1:2000SEP08	g4569925	421	731
136	LG:1087811.1:2000SEP08	g2706016	427	727
136	LG:1087811.1:2000SEP08	5294637H1	432	670
136	LG:1087811.1:2000SEP08	6350519H2	437	718
136	LG:1087811.1:2000SEP08	6743390H1	438	942
136	LG:1087811.1:2000SEP08	5263067H1	459	547
136	LG:1087811.1:2000SEP08	3800808H1	502	687
136	LG:1087811.1:2000SEP08	g991148	527	728
136	LG:1087811.1:2000SEP08	6252361F8	13	539
136	LG:1087811.1:2000SEP08	6521954H1	13	310
136	LG:1087811.1:2000SEP08	450654H1	13	117
136	LG:1087811.1:2000SEP08	7655877J1	13	64
136	LG:1087811.1:2000SEP08	339857H1	13	121
136	LG:1087811.1:2000SEP08	1377096H1	13	128
136	LG:1087811.1:2000SEP08	4951185H1	13	162
136	LG:1087811.1:2000SEP08	5960651H1	13	271
136	LG:1087811.1:2000SEP08	7703548J1	13	509
136	LG:1087811.1:2000SEP08	6015313H1	13	268
137	LG:1327885.1:2000SEP08	6798095T8	1	537
137	LG:1327885.1:2000SEP08	6798095H1	1	554

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
137	LG:1327885.1:2000SEP08	6798095F8	1	519
138	LI:449393.1:2000SEP08	6272292H2	1	507
138	LI:449393.1:2000SEP08	6272292F8	1	642
138	LI:449393.1:2000SEP08	5910821F8	365	793
138	LI:449393.1:2000SEP08	5910821T9	365	662
138	LI:449393.1:2000SEP08	5910821T8	365	636
138	LI:449393.1:2000SEP08	5910821H1	365	676
139	LI:897616.1:2000SEP08	6790830H1	1	523
139	LI:897616.1:2000SEP08	3649045H1	1	143
139	LI:897616.1:2000SEP08	6796674H1	2	519
139	LI:897616.1:2000SEP08	6790830T8	1	498
139	LI:897616.1:2000SEP08	6796674T8	2	442
139	LI:897616.1:2000SEP08	6796674F8	2	525
139	LI:897616.1:2000SEP08	6796548T8	19	470
139	LI:897616.1:2000SEP08	6796548H1	21	449
139	LI:897616.1:2000SEP08	6796548F8	21	526
139	LI:897616.1:2000SEP08	6790830F8	30	539
140	LI:736860.1:2000SEP08	6274719H2	1	274
140	LI:736860.1:2000SEP08	6274719T8	1	331
140	LI:736860.1:2000SEP08	6274719F8	15	431
141	LI:027066.6:2000SEP08	2101767R6	1	437
141	LI:027066.6:2000SEP08	6828854H1	307	892
141	LI:027066.6:2000SEP08	7652140H1	645	947
141	LI:027066.6:2000SEP08	6080233H1	640	1139
141	LI:027066.6:2000SEP08	7194730H1	641	1114
141	LI:027066.6:2000SEP08	6827816J1	843	962
141	LI:027066.6:2000SEP08	2593628F6	855	1205
141	LI:027066.6:2000SEP08	2593628H1	855	1102
141	LI:027066.6:2000SEP08	4725176H1	870	1111
141	LI:027066.6:2000SEP08	3291113H1	910	1170
141	LI:027066.6:2000SEP08	5328747H1	944	1097
141	LI:027066.6:2000SEP08	3599076H1	983	1254
141	LI:027066.6:2000SEP08	1821378H1	1015	1120
141	LI:027066.6:2000SEP08	3013226H1	1029	1326
141	LI:027066.6:2000SEP08	2776582H1	1176	1429
142	LI:1074263.1:2000SEP08	6792763H1	1	460
142	LI:1074263.1:2000SEP08	6792018T8	1	567
142	LI:1074263.1:2000SEP08	6792018F8	1	649
142	LI:1074263.1:2000SEP08	6792018H1	1	406
142	LI:1074263.1:2000SEP08	6796374H1	11	503
142	LI:1074263.1:2000SEP08	6791238H1	41	618
142	LI:1074263.1:2000SEP08	6791238F8	41	637
142	LI:1074263.1:2000SEP08	6795926H1	42	461
142	LI:1074263.1:2000SEP08	6791238T8	55	595
142	LI:1074263.1:2000SEP08	6792967H1	46	595
142	LI:1074263.1:2000SEP08	6794141H1	152	652
142	LI:1074263.1:2000SEP08	6796228H1	145	654
142	LI:1074263.1:2000SEP08	6794232H1	145	661
142	LI:1074263.1:2000SEP08	6798243H1	147	663

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
142	LI:1074263.1:2000SEP08	6794141F8	152	620
142	LI:1074263.1:2000SEP08	6791574H1	149	663
142	LI:1074263.1:2000SEP08	6795827H1	347	661
142	LI:1074263.1:2000SEP08	6794232T8	468	529
143	LI:334345.1:2000SEP08	2108692T6	100	701
143	LI:334345.1:2000SEP08	2108692R6	75	558
143	LI:334345.1:2000SEP08	2108692H1	75	335
143	LI:334345.1:2000SEP08	60210386U1	253	458
143	LI:334345.1:2000SEP08	60220684D1	1	328
144	LI:1093914.1:2000SEP08	6867381H1	349	897
144	LI:1093914.1:2000SEP08	6867274H1	348	893
144	LI:1093914.1:2000SEP08	6867281H1	348	889
144	LI:1093914.1:2000SEP08	6867373H1	347	857
144	LI:1093914.1:2000SEP08	g3094405	702	793
144	LI:1093914.1:2000SEP08	g2702894	1	437
144	LI:1093914.1:2000SEP08	g3959241	49	436
144	LI:1093914.1:2000SEP08	g712168	206	439
144	LI:1093914.1:2000SEP08	g5113968	714	920
144	LI:1093914.1:2000SEP08	g3959109	43	436
144	LI:1093914.1:2000SEP08	g3279579	48	437
145	LI:1188168.1:2000SEP08	6621226J1	901	1279
145	LI:1188168.1:2000SEP08	3746627F6	928	1283
145	LI:1188168.1:2000SEP08	3908920H1	911	1172
145	LI:1188168.1:2000SEP08	70185233V1	942	1336
145	LI:1188168.1:2000SEP08	3746627H1	928	1210
145	LI:1188168.1:2000SEP08	5672277H1	1991	2237
145	LI:1188168.1:2000SEP08	472727H1	2017	2325
145	LI:1188168.1:2000SEP08	6569428H1	2015	2234
145	LI:1188168.1:2000SEP08	1622655H1	2534	2753
145	LI:1188168.1:2000SEP08	g2218626	2530	2753
145	LI:1188168.1:2000SEP08	70184129V1	1845	2447
145	LI:1188168.1:2000SEP08	70594080V1	1848	2417
145	LI:1188168.1:2000SEP08	1858113F6	1846	2326
145	LI:1188168.1:2000SEP08	70592770V1	1856	2256
145	LI:1188168.1:2000SEP08	70184080V1	1632	2009
145	LI:1188168.1:2000SEP08	7104537R8	2158	2734
145	LI:1188168.1:2000SEP08	6404622T8	2163	2639
145	LI:1188168.1:2000SEP08	6120976H1	1793	2072
145	LI:1188168.1:2000SEP08	6875307H1	2343	2753
145	LI:1188168.1:2000SEP08	70185768V1	473	788
145	LI:1188168.1:2000SEP08	70185230V1	485	850
145	LI:1188168.1:2000SEP08	6983079H1	542	1045
145	LI:1188168.1:2000SEP08	4289648H1	550	802
145	LI:1188168.1:2000SEP08	2594169H1	1985	2265
145	LI:1188168.1:2000SEP08	6033282H1	740	1364
145	LI:1188168.1:2000SEP08	1670211H1	754	882
145	LI:1188168.1:2000SEP08	1667958H1	754	866
145	LI:1188168.1:2000SEP08	70185618V1	753	1138
145	LI:1188168.1:2000SEP08	6900769H1	798	1227

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
145	LI:1188168.1:2000SEP08	7583166H1	1	463
145	LI:1188168.1:2000SEP08	3746627T6	2360	2716
145	LI:1188168.1:2000SEP08	70613203V1	1446	1628
145	LI:1188168.1:2000SEP08	70184608V1	702	1175
145	LI:1188168.1:2000SEP08	70184435V1	1494	2040
145	LI:1188168.1:2000SEP08	6437826H1	1494	1966
145	LI:1188168.1:2000SEP08	2110610H1	1508	1791
145	LI:1188168.1:2000SEP08	6893096H1	2130	2729
145	LI:1188168.1:2000SEP08	367946H1	2128	2273
145	LI:1188168.1:2000SEP08	70185854V1	1638	2097
145	LI:1188168.1:2000SEP08	5530038H1	1662	1880
145	LI:1188168.1:2000SEP08	70594169V1	1846	2468
145	LI:1188168.1:2000SEP08	70591460V1	1846	2564
145	LI:1188168.1:2000SEP08	g4736011	2292	2753
145	LI:1188168.1:2000SEP08	g3277795	2290	2756
145	LI:1188168.1:2000SEP08	g4896720	2314	2748
145	LI:1188168.1:2000SEP08	6119621H1	1793	2415
145	LI:1188168.1:2000SEP08	70184766V1	554	897
145	LI:1188168.1:2000SEP08	70185614V1	626	941
145	LI:1188168.1:2000SEP08	70184328V1	645	1095
145	LI:1188168.1:2000SEP08	70184692V1	663	1077
145	LI:1188168.1:2000SEP08	8048532H1	711	1404
145	LI:1188168.1:2000SEP08	4557603H1	2014	2195
145	LI:1188168.1:2000SEP08	7417590T1	2040	2646
145	LI:1188168.1:2000SEP08	70185294V1	1167	1480
145	LI:1188168.1:2000SEP08	7710808H1	1181	1468
145	LI:1188168.1:2000SEP08	70186349V1	1182	1710
145	LI:1188168.1:2000SEP08	70186009V1	1241	1660
145	LI:1188168.1:2000SEP08	70593641V1	1846	2351
145	LI:1188168.1:2000SEP08	70184639V1	196	642
145	LI:1188168.1:2000SEP08	70185751V1	196	490
145	LI:1188168.1:2000SEP08	2914058H1	2662	2744
145	LI:1188168.1:2000SEP08	6209724H1	1783	1964
145	LI:1188168.1:2000SEP08	70595344V1	1845	2487
145	LI:1188168.1:2000SEP08	70592046V1	1856	2296
145	LI:1188168.1:2000SEP08	70184357V1	1844	2451
145	LI:1188168.1:2000SEP08	2909961H1	17	283
145	LI:1188168.1:2000SEP08	3216652F6	196	474
145	LI:1188168.1:2000SEP08	70185750V1	196	251
145	LI:1188168.1:2000SEP08	g3240143	2315	2757
145	LI:1188168.1:2000SEP08	70184923V1	1386	1648
145	LI:1188168.1:2000SEP08	3121125H1	1392	1725
145	LI:1188168.1:2000SEP08	70186237V1	1406	1814
145	LI:1188168.1:2000SEP08	1418626H1	1439	1686
145	LI:1188168.1:2000SEP08	6893096J1	1414	1992
145	LI:1188168.1:2000SEP08	70186177V1	1420	1810
145	LI:1188168.1:2000SEP08	4001227H1	2348	2666
145	LI:1188168.1:2000SEP08	537152H1	2141	2402
145	LI:1188168.1:2000SEP08	1902576T6	2149	2711



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
145	LI:1188168.1:2000SEP08	2951572T7	2145	2707
145	LI:1188168.1:2000SEP08	3042756H1	964	1243
145	LI:1188168.1:2000SEP08	70185592V1	989	1380
145	LI:1188168.1:2000SEP08	70184831V1	951	1497
145	LI:1188168.1:2000SEP08	70185438V1	1053	1529
145	LI:1188168.1:2000SEP08	2951572H1	1824	2138
145	LI:1188168.1:2000SEP08	5389255H1	1822	1971
145	LI:1188168.1:2000SEP08	2951572F7	1824	2396
145	LI:1188168.1:2000SEP08	5844552H1	1737	1815
145	LI:1188168.1:2000SEP08	6182216H1	1765	1947
145	LI:1188168.1:2000SEP08	70186007V1	1767	2173
145	LI:1188168.1:2000SEP08	6030444H1	1817	2123
145	LI:1188168.1:2000SEP08	6058238H1	1817	1968
145	LI:1188168.1:2000SEP08	g2253977	2278	2755
145	LI:1188168.1:2000SEP08	7436362H1	344	963
145	LI:1188168.1:2000SEP08	7678852H1	441	1040
145	LI:1188168.1:2000SEP08	70186262V1	434	914
145	LI:1188168.1:2000SEP08	70185594V1	468	919
145	LI:1188168.1:2000SEP08	70599167V1	1445	1550
145	LI:1188168.1:2000SEP08	4950546H1	2113	2406
145	LI:1188168.1:2000SEP08	70574741V1	2045	2361
145	LI:1188168.1:2000SEP08	4367125H1	2287	2572
145	LI:1188168.1:2000SEP08	70184650V1	1461	2062
145	LI:1188168.1:2000SEP08	70184019V1	1845	2517
145	LI:1188168.1:2000SEP08	70184585V1	1619	2209
145	LI:1188168.1:2000SEP08	1622843H1	2534	2753
145	LI:1188168.1:2000SEP08	037358H1	2117	2305
145	LI:1188168.1:2000SEP08	2911254H1	1525	1802
145	LI:1188168.1:2000SEP08	6814276J1	1546	2165
145	LI:1188168.1:2000SEP08	70594137V1	1846	2326
145	LI:1188168.1:2000SEP08	4462757H1	2245	2510
145	LI:1188168.1:2000SEP08	70592617V1	2249	2784
145	LI:1188168.1:2000SEP08	3617928H1	2266	2595
145	LI:1188168.1:2000SEP08	1858113T6	2271	2714
145	LI:1188168.1:2000SEP08	6167723H1	2378	2744
145	LI:1188168.1:2000SEP08	1964481T6	2379	2706
145	LI:1188168.1:2000SEP08	3216652H2	196	450
145	LI:1188168.1:2000SEP08	6946538H1	216	711
145	LI:1188168.1:2000SEP08	70184645V1	230	709
145	LI:1188168.1:2000SEP08	7038575H1	281	537
145	LI:1188168.1:2000SEP08	7398782H2	120	706
145	LI:1188168.1:2000SEP08	g1750964	2587	2744
145	LI:1188168.1:2000SEP08	7289561H1	1826	2189
145	LI:1188168.1:2000SEP08	70595411V1	1858	2365
145	LI:1188168.1:2000SEP08	70186306V1	1857	2297
145	LI:1188168.1:2000SEP08	2821727H1	1850	2192
145	LI:1188168.1:2000SEP08	70184672V1	726	1270
145	LI:1188168.1:2000SEP08	70592744V1	1847	2146
145	LI:1188168.1:2000SEP08	70596593V1	1846	2442

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
145	LI:1188168.1:2000SEP08	70599937V1	1682	2268
145	LI:1188168.1:2000SEP08	70185955V1	1728	2141
145	LI:1188168.1:2000SEP08	3289040T6	2387	2706
145	LI:1188168.1:2000SEP08	1964481R6	2403	2747
145	LI:1188168.1:2000SEP08	1964473H1	2403	2468
145	LI:1188168.1:2000SEP08	70185906V1	2380	2751
145	LI:1188168.1:2000SEP08	1964481H1	2403	2457
145	LI:1188168.1:2000SEP08	6951742H1	2421	2670
145	LI:1188168.1:2000SEP08	g4876419	2425	2754
145	LI:1188168.1:2000SEP08	70185796V1	798	1088
145	LI:1188168.1:2000SEP08	70185361V1	816	1309
145	LI:1188168.1:2000SEP08	4580748H1	860	1016
145	LI:1188168.1:2000SEP08	4844929H2	2446	2718
145	LI:1188168.1:2000SEP08	70596434V1	1951	2600
145	LI:1188168.1:2000SEP08	643662H1	1950	2260
145	LI:1188168.1:2000SEP08	6264871H1	1967	2615
145	LI:1188168.1:2000SEP08	5107681H1	1940	2108
145	LI:1188168.1:2000SEP08	70186335V1	1615	2208
145	LI:1188168.1:2000SEP08	7614738H1	1807	2483
145	LI:1188168.1:2000SEP08	70185166V1	1322	1700
145	LI:1188168.1:2000SEP08	5525920H2	1322	1566
145	LI:1188168.1:2000SEP08	2119738R6	1324	1767
145	LI:1188168.1:2000SEP08	2119738H1	1324	1586
145	LI:1188168.1:2000SEP08	70184709V1	1328	1778
145	LI:1188168.1:2000SEP08	70184773V1	1351	1778
145	LI:1188168.1:2000SEP08	2570715H1	1358	1611
145	LI:1188168.1:2000SEP08	70185090V1	1362	1887
145	LI:1188168.1:2000SEP08	70185804V1	1374	1717
145	LI:1188168.1:2000SEP08	6891322J1	1383	2067
145	LI:1188168.1:2000SEP08	g775086	1381	1627
145	LI:1188168.1:2000SEP08	70184416V1	1632	2327
145	LI:1188168.1:2000SEP08	5286785H1	1572	1742
145	LI:1188168.1:2000SEP08	70186202V1	1583	2035
145	LI:1188168.1:2000SEP08	1902576H1	1445	1713
145	LI:1188168.1:2000SEP08	70598884V1	1445	1926
145	LI:1188168.1:2000SEP08	1902576F6	1445	1962
145	LI:1188168.1:2000SEP08	70602407V1	1445	1988
145	LI:1188168.1:2000SEP08	1609475T6	2149	2713
145	LI:1188168.1:2000SEP08	7410153H1	1054	1550
145	LI:1188168.1:2000SEP08	4135591H1	1061	1355
145	LI:1188168.1:2000SEP08	70185778V1	1068	1449
145	LI:1188168.1:2000SEP08	70185190V1	1147	1476
145	LI:1188168.1:2000SEP08	70185325V1	1164	1685
145	LI:1188168.1:2000SEP08	70593353V1	2293	2754
145	LI:1188168.1:2000SEP08	70593900V1	1858	2008
145	LI:1188168.1:2000SEP08	70184704V1	1859	2364
145	LI:1188168.1:2000SEP08	70594391V1	1868	2522
145	LI:1188168.1:2000SEP08	3165818H1	1881	2189
145	LI:1188168.1:2000SEP08	1858113H1	1846	2148

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
145	U:1188168.1:2000SEP08	70185264V1	1849	2108
145	U:1188168.1:2000SEP08	70184698V1	1895	2398
145	U:1188168.1:2000SEP08	1609475F6	1911	2398
145	U:1188168.1:2000SEP08	1609475H1	1911	2144
145	U:1188168.1:2000SEP08	6334570H1	1970	2434
145	U:1188168.1:2000SEP08	4510027H1	1981	2294
145	U:1188168.1:2000SEP08	g2466512	2456	2649
145	U:1188168.1:2000SEP08	2622543H1	2449	2729
145	U:1188168.1:2000SEP08	3488589H1	1248	1512
145	U:1188168.1:2000SEP08	70186221V1	1293	1941
145	U:1188168.1:2000SEP08	70185993V1	1298	1540
145	U:1188168.1:2000SEP08	70185509V1	1322	1792
145	U:1188168.1:2000SEP08	g2877323	2461	2753
145	U:1188168.1:2000SEP08	g822020	2512	2762
145	U:1188168.1:2000SEP08	7594855H1	2527	2753
145	U:1188168.1:2000SEP08	1609420H1	1911	2105
145	U:1188168.1:2000SEP08	3751812H1	1909	2149
145	U:1188168.1:2000SEP08	2600360H1	1909	2164
145	U:1188168.1:2000SEP08	2836303H1	1928	2223
145	U:1188168.1:2000SEP08	70595428V1	1926	2300
145	U:1188168.1:2000SEP08	2119738T6	2167	2697
145	U:1188168.1:2000SEP08	6206776H1	2182	2612
145	U:1188168.1:2000SEP08	70596229V1	2182	2381
145	U:1188168.1:2000SEP08	4087496H1	2197	2503
145	U:1188168.1:2000SEP08	7415691T2	2224	2644
145	U:1188168.1:2000SEP08	70185549V1	2223	2665
145	U:1188168.1:2000SEP08	2614105T6	2226	2697
145	U:1188168.1:2000SEP08	6016239H1	2236	2713
145	U:1188168.1:2000SEP08	70185605V1	2164	2659
145	U:1188168.1:2000SEP08	70593879V1	2165	2532
145	U:1188168.1:2000SEP08	70185613V1	2164	2602
145	U:1188168.1:2000SEP08	70185958V1	1637	2096
145	U:1188168.1:2000SEP08	70601381V1	2117	2570
145	U:1188168.1:2000SEP08	70596469V1	2119	2320
145	U:1188168.1:2000SEP08	1686731H1	2703	2757
145	U:1188168.1:2000SEP08	g2148494	2670	2753
146	U:1065168.1:2000SEP08	6795315F8	1	630
146	U:1065168.1:2000SEP08	6795315H1	1	69
146	U:1065168.1:2000SEP08	6795315T8	143	584
147	U:1180418.1:2000SEP08	71600275V1	464	620
147	U:1180418.1:2000SEP08	3800808H1	502	687
147	U:1180418.1:2000SEP08	g991148	527	728
147	U:1180418.1:2000SEP08	2304130H1	539	798
147	U:1180418.1:2000SEP08	g5362427	542	998
147	U:1180418.1:2000SEP08	6387991H1	555	718
147	U:1180418.1:2000SEP08	6590088H1	626	718
147	U:1180418.1:2000SEP08	5043837H1	666	930
147	U:1180418.1:2000SEP08	g5744437	761	1003
147	U:1180418.1:2000SEP08	71598470V1	771	1306

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
147	LI:1180418.1:2000SEP08	1223772H1	784	1033
147	LI:1180418.1:2000SEP08	853227T6	803	1019
147	LI:1180418.1:2000SEP08	3624479H1	815	1019
147	LI:1180418.1:2000SEP08	7658126J1	826	1338
147	LI:1180418.1:2000SEP08	1942610H1	845	1015
147	LI:1180418.1:2000SEP08	1942610R6	845	1019
147	LI:1180418.1:2000SEP08	1942610T6	845	1019
147	LI:1180418.1:2000SEP08	1447494H1	849	1003
147	LI:1180418.1:2000SEP08	144709F1	878	1019
147	LI:1180418.1:2000SEP08	146431H1	883	1072
147	LI:1180418.1:2000SEP08	144709H1	889	1015
147	LI:1180418.1:2000SEP08	144709R1	890	1019
147	LI:1180418.1:2000SEP08	g2264487	952	1019
147	LI:1180418.1:2000SEP08	71601007V1	764	1019
147	LI:1180418.1:2000SEP08	342750H1	141	362
147	LI:1180418.1:2000SEP08	4441395H1	165	281
147	LI:1180418.1:2000SEP08	376238H1	187	432
147	LI:1180418.1:2000SEP08	2352459H1	213	450
147	LI:1180418.1:2000SEP08	2351409H1	213	443
147	LI:1180418.1:2000SEP08	565688F1	214	728
147	LI:1180418.1:2000SEP08	2766321H1	217	459
147	LI:1180418.1:2000SEP08	6484730H1	236	644
147	LI:1180418.1:2000SEP08	501253H1	236	396
147	LI:1180418.1:2000SEP08	518551F1	241	728
147	LI:1180418.1:2000SEP08	4878173H1	243	507
147	LI:1180418.1:2000SEP08	1746545H1	243	511
147	LI:1180418.1:2000SEP08	497017H1	256	448
147	LI:1180418.1:2000SEP08	3121516H1	257	542
147	LI:1180418.1:2000SEP08	g5837313	275	727
147	LI:1180418.1:2000SEP08	g2704490	278	728
147	LI:1180418.1:2000SEP08	6725179F8	298	852
147	LI:1180418.1:2000SEP08	498297H1	299	436
147	LI:1180418.1:2000SEP08	g1099349	299	728
147	LI:1180418.1:2000SEP08	g2279564	301	730
147	LI:1180418.1:2000SEP08	g6398356	301	728
147	LI:1180418.1:2000SEP08	6725179H1	298	852
147	LI:1180418.1:2000SEP08	g5395790	305	731
147	LI:1180418.1:2000SEP08	g7150021	326	728
147	LI:1180418.1:2000SEP08	g6657428	326	718
147	LI:1180418.1:2000SEP08	341786H1	331	530
147	LI:1180418.1:2000SEP08	5537161H1	350	544
147	LI:1180418.1:2000SEP08	71600737V1	360	807
147	LI:1180418.1:2000SEP08	478130H1	370	643
147	LI:1180418.1:2000SEP08	565688T6	364	684
147	LI:1180418.1:2000SEP08	1221285H1	392	623
147	LI:1180418.1:2000SEP08	7958108J1	407	1065
147	LI:1180418.1:2000SEP08	g2953838	407	731
147	LI:1180418.1:2000SEP08	563428H1	421	650
147	LI:1180418.1:2000SEP08	7658126H1	420	901

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
147	LI:1180418.1:2000SEP08	853227R6	68	574
147	LI:1180418.1:2000SEP08	518551R1	74	547
147	LI:1180418.1:2000SEP08	7740966H1	79	722
147	LI:1180418.1:2000SEP08	7958108H1	99	725
147	LI:1180418.1:2000SEP08	339831H1	105	341
147	LI:1180418.1:2000SEP08	500238F1	128	728
147	LI:1180418.1:2000SEP08	g4569925	421	731
147	LI:1180418.1:2000SEP08	g2706016	427	727
147	LI:1180418.1:2000SEP08	5294637H1	432	670
147	LI:1180418.1:2000SEP08	6350519H2	437	718
147	LI:1180418.1:2000SEP08	6743390H1	438	942
147	LI:1180418.1:2000SEP08	71599677V1	490	1019
147	LI:1180418.1:2000SEP08	6252269H1	1	317
147	LI:1180418.1:2000SEP08	6252361H1	1	249
147	LI:1180418.1:2000SEP08	g758862	1	170
147	LI:1180418.1:2000SEP08	939450H1	1	138
147	LI:1180418.1:2000SEP08	1875744H1	1	254
147	LI:1180418.1:2000SEP08	1873658H1	1	273
147	LI:1180418.1:2000SEP08	345602H1	1	222
147	LI:1180418.1:2000SEP08	3563357H1	6	256
147	LI:1180418.1:2000SEP08	565688R1	7	459
147	LI:1180418.1:2000SEP08	565688R6	7	291
147	LI:1180418.1:2000SEP08	5199726H1	7	270
147	LI:1180418.1:2000SEP08	565688H1	7	229
147	LI:1180418.1:2000SEP08	7724194J1	13	454
147	LI:1180418.1:2000SEP08	6167719H1	13	491
147	LI:1180418.1:2000SEP08	7655877J1	13	64
147	LI:1180418.1:2000SEP08	3396040H1	13	68
147	LI:1180418.1:2000SEP08	7655877H1	13	67
147	LI:1180418.1:2000SEP08	338433H1	13	82
147	LI:1180418.1:2000SEP08	4951185H1	13	162
147	LI:1180418.1:2000SEP08	6521954H1	13	310
147	LI:1180418.1:2000SEP08	450654H1	13	117
147	LI:1180418.1:2000SEP08	6252361F8	13	539
147	LI:1180418.1:2000SEP08	339857H1	13	121
147	LI:1180418.1:2000SEP08	7724194H1	13	615
147	LI:1180418.1:2000SEP08	1377096H1	13	128
147	LI:1180418.1:2000SEP08	6015313H1	13	268
147	LI:1180418.1:2000SEP08	5960651H1	13	271
147	LI:1180418.1:2000SEP08	7703548J1	13	509
147	LI:1180418.1:2000SEP08	5340183F8	15	247
147	LI:1180418.1:2000SEP08	6252361T8	19	629
147	LI:1180418.1:2000SEP08	4803270H1	21	290
147	LI:1180418.1:2000SEP08	564726H1	22	241
147	LI:1180418.1:2000SEP08	500238H1	22	227
147	LI:1180418.1:2000SEP08	500238R1	22	421
147	LI:1180418.1:2000SEP08	5805876H1	28	235
147	LI:1180418.1:2000SEP08	375720H1	27	292
147	LI:1180418.1:2000SEP08	518551H1	50	190

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
147	LI:1180418.1:2000SEP08	3489279H1	51	315
147	LI:1180418.1:2000SEP08	5340183T8	58	630
147	LI:1180418.1:2000SEP08	4848110H1	65	154
147	LI:1180418.1:2000SEP08	853227H1	68	315
147	LI:1180418.1:2000SEP08	71599552V1	779	1019
147	LI:1180418.1:2000SEP08	71602555V1	788	1019
147	LI:1180418.1:2000SEP08	71602383V1	830	1019
148	LG:232648.1:2000SEP08	3437327H1	126	312
148	LG:232648.1:2000SEP08	2471325H1	124	353
148	LG:232648.1:2000SEP08	1665878H1	119	357
148	LG:232648.1:2000SEP08	7679565H1	1	550
148	LG:232648.1:2000SEP08	6038768H1	105	319
148	LG:232648.1:2000SEP08	2625304H1	474	733
148	LG:232648.1:2000SEP08	6804355R8	379	543
148	LG:232648.1:2000SEP08	6804355F8	379	542
148	LG:232648.1:2000SEP08	6804355H1	379	543
148	LG:232648.1:2000SEP08	6804355J1	362	543
148	LG:232648.1:2000SEP08	3742927H1	126	415
148	LG:232648.1:2000SEP08	6038768F8	129	724
148	LG:232648.1:2000SEP08	2464965H1	130	363
148	LG:232648.1:2000SEP08	3856515F8	321	894
148	LG:232648.1:2000SEP08	g2556959	1815	2142
148	LG:232648.1:2000SEP08	484034R1	1768	2142
148	LG:232648.1:2000SEP08	484034T7	1770	2063
148	LG:232648.1:2000SEP08	g5447846	1781	2141
148	LG:232648.1:2000SEP08	7700719H1	1782	2134
148	LG:232648.1:2000SEP08	1619047H1	1765	1982
148	LG:232648.1:2000SEP08	6213262H1	1767	2049
148	LG:232648.1:2000SEP08	g3330066	1769	2145
148	LG:232648.1:2000SEP08	7239656H1	1666	2125
148	LG:232648.1:2000SEP08	7056066H1	1674	2142
148	LG:232648.1:2000SEP08	484034F1	1766	2082
148	LG:232648.1:2000SEP08	6127977H1	1697	2152
148	LG:232648.1:2000SEP08	g4902141	1721	2130
148	LG:232648.1:2000SEP08	7450786T1	1733	2010
148	LG:232648.1:2000SEP08	4505914H1	1733	1992
148	LG:232648.1:2000SEP08	5903286H1	1733	2005
148	LG:232648.1:2000SEP08	4664636H1	1748	2000
148	LG:232648.1:2000SEP08	4664658H1	1748	2001
148	LG:232648.1:2000SEP08	g885174	1754	2162
148	LG:232648.1:2000SEP08	g2726406	1753	2095
148	LG:232648.1:2000SEP08	3739527H1	1850	2141
148	LG:232648.1:2000SEP08	g4989863	1868	2142
148	LG:232648.1:2000SEP08	g2009754	1846	2089
148	LG:232648.1:2000SEP08	3739527F7	1849	2142
148	LG:232648.1:2000SEP08	6017442H1	1475	2066
148	LG:232648.1:2000SEP08	2951610H1	1825	2111
148	LG:232648.1:2000SEP08	632399R6	2015	2152
148	LG:232648.1:2000SEP08	632399T6	2015	2093

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
148	LG:232648.1:2000SEP08	602162T6	1926	2102
148	LG:232648.1:2000SEP08	602162R6	1926	2145
148	LG:232648.1:2000SEP08	484034H1	1960	2142
148	LG:232648.1:2000SEP08	7941209H1	1973	2145
148	LG:232648.1:2000SEP08	g2224223	1885	2152
148	LG:232648.1:2000SEP08	g4150209	2042	2148
148	LG:232648.1:2000SEP08	632399H1	2015	2150
148	LG:232648.1:2000SEP08	g2398138	2020	2145
148	LG:232648.1:2000SEP08	g5672245	2029	2142
148	LG:232648.1:2000SEP08	g3425208	2029	2153
148	LG:232648.1:2000SEP08	6850278H1	1470	2022
148	LG:232648.1:2000SEP08	3856515T8	1467	2042
148	LG:232648.1:2000SEP08	6544359H1	1464	2015
148	LG:232648.1:2000SEP08	586344H1	1536	1803
148	LG:232648.1:2000SEP08	g896529	1540	1811
148	LG:232648.1:2000SEP08	2740488H1	1505	1741
148	LG:232648.1:2000SEP08	1292945H1	1482	1742
148	LG:232648.1:2000SEP08	6038768T8	1488	2041
148	LG:232648.1:2000SEP08	1292945F6	1482	1858
148	LG:232648.1:2000SEP08	3397313H1	1452	1699
148	LG:232648.1:2000SEP08	1665878T6	1656	2104
148	LG:232648.1:2000SEP08	2781935H1	1656	1900
148	LG:232648.1:2000SEP08	5094116H1	1578	1839
148	LG:232648.1:2000SEP08	1292945T6	1642	2104
148	LG:232648.1:2000SEP08	5307388H1	1542	1772
148	LG:232648.1:2000SEP08	5306755H1	1542	1742
148	LG:232648.1:2000SEP08	4857661H1	1547	1685
148	LG:232648.1:2000SEP08	5161538H1	1551	1781
148	LG:232648.1:2000SEP08	5306855H1	1542	1658
148	LG:232648.1:2000SEP08	1932006F6	563	990
148	LG:232648.1:2000SEP08	g774443	670	890
148	LG:232648.1:2000SEP08	3083767H1	747	998
148	LG:232648.1:2000SEP08	3616968H1	839	1082
148	LG:232648.1:2000SEP08	522198H1	1013	1244
148	LG:232648.1:2000SEP08	7679565J1	489	1100
148	LG:232648.1:2000SEP08	963051R2	521	825
148	LG:232648.1:2000SEP08	963051H1	521	857
148	LG:232648.1:2000SEP08	7234232H1	550	955
148	LG:232648.1:2000SEP08	1932006H1	563	823
148	LG:232648.1:2000SEP08	5902804H1	1419	1643
148	LG:232648.1:2000SEP08	g1962798	1223	1382
148	LG:232648.1:2000SEP08	6161987H1	1444	1946
148	LG:232648.1:2000SEP08	820544H1	1222	1371
148	LG:232648.1:2000SEP08	820544R1	1222	1771
148	LG:232648.1:2000SEP08	6429960H1	1269	1595
148	LG:232648.1:2000SEP08	g1548524	1291	1786
148	LG:232648.1:2000SEP08	6409782H1	1372	1651
148	LG:232648.1:2000SEP08	1535447H1	1199	1405
148	LG:232648.1:2000SEP08	6572961H1	1201	1706

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
148	LG:232648.1:2000SEP08	6077657H1	1211	1283
148	LG:232648.1:2000SEP08	7700719J1	1064	1673
148	LG:232648.1:2000SEP08	3948553H1	1072	1347
148	LG:232648.1:2000SEP08	6291826H1	1102	1240
148	LG:232648.1:2000SEP08	6149863H1	1119	1540
148	LG:232648.1:2000SEP08	g888070	1184	1444
149	LG:1078420.1:2000SEP08	3199584F6	5	508
149	LG:1078420.1:2000SEP08	2617918F6	120	367
149	LG:1078420.1:2000SEP08	3004192F6	5	578
149	LG:1078420.1:2000SEP08	2966808F6	5	585
149	LG:1078420.1:2000SEP08	3051156F6	5	593
149	LG:1078420.1:2000SEP08	4331527F6	1	586
149	LG:1078420.1:2000SEP08	4331527H1	1	140
149	LG:1078420.1:2000SEP08	2929058F6	5	594
149	LG:1078420.1:2000SEP08	2990863F6	5	528
149	LG:1078420.1:2000SEP08	2929058T6	522	972
149	LG:1078420.1:2000SEP08	3076784F6	5	555
149	LG:1078420.1:2000SEP08	g3428455	17	505
149	LG:1078420.1:2000SEP08	g3674918	17	479
149	LG:1078420.1:2000SEP08	3001577F6	5	585
150	LG:1397599.1:2000SEP08	g678960	481	669
150	LG:1397599.1:2000SEP08	7082955H1	559	696
150	LG:1397599.1:2000SEP08	4625011T8	42	628
150	LG:1397599.1:2000SEP08	2807456F6	1	506
150	LG:1397599.1:2000SEP08	4625011F9	1	506
150	LG:1397599.1:2000SEP08	7630459J1	1	474
150	LG:1397599.1:2000SEP08	6750082H1	202	309
150	LG:1397599.1:2000SEP08	2807456H1	1	247
150	LG:1397599.1:2000SEP08	2807456T6	120	669
150	LG:1397599.1:2000SEP08	7660359H1	71	660
151	LG:1397655.2:2000SEP08	587588R6	414	749
151	LG:1397655.2:2000SEP08	7382264H1	338	738
151	LG:1397655.2:2000SEP08	7237485H1	1	486
151	LG:1397655.2:2000SEP08	g2806538	579	969
151	LG:1397655.2:2000SEP08	g3237896	643	968
151	LG:1397655.2:2000SEP08	g4149271	782	968
151	LG:1397655.2:2000SEP08	7658736J1	338	819
151	LG:1397655.2:2000SEP08	2839106T6	356	937
151	LG:1397655.2:2000SEP08	587588T6	414	925
151	LG:1397655.2:2000SEP08	2406278H1	660	890
151	LG:1397655.2:2000SEP08	g3179231	631	977
151	LG:1397655.2:2000SEP08	g3110195	508	971
152	LG:241055.1:2000SEP08	2600825F6	12	305
152	LG:241055.1:2000SEP08	2662733H1	1	224
152	LG:241055.1:2000SEP08	5732381F6	117	580
152	LG:241055.1:2000SEP08	2600825H1	9	89
152	LG:241055.1:2000SEP08	g2809760	227	688
152	LG:241055.1:2000SEP08	6137035H1	253	558
152	LG:241055.1:2000SEP08	4302282H1	255	537



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
152	LG:241055.1:2000SEP08	2617918H1	257	469
152	LG:241055.1:2000SEP08	1292452H1	291	551
152	LG:241055.1:2000SEP08	g5368192	225	691
152	LG:241055.1:2000SEP08	4714957H1	141	394
152	LG:241055.1:2000SEP08	3199584H1	142	418
152	LG:241055.1:2000SEP08	4163185H1	171	454
152	LG:241055.1:2000SEP08	6259513H1	206	491
152	LG:241055.1:2000SEP08	g5368211	225	675
152	LG:241055.1:2000SEP08	7048864H1	581	1152
152	LG:241055.1:2000SEP08	2830919H1	1049	1265
152	LG:241055.1:2000SEP08	409881H1	1101	1265
152	LG:241055.1:2000SEP08	g2934256	299	617
152	LG:241055.1:2000SEP08	405951H1	358	602
152	LG:241055.1:2000SEP08	2435520H1	411	627
152	LG:241055.1:2000SEP08	6162384H1	447	658
152	LG:241055.1:2000SEP08	2785236H2	451	726
152	LG:241055.1:2000SEP08	1235959H1	498	660
152	LG:241055.1:2000SEP08	7131488H1	581	1137
152	LG:241055.1:2000SEP08	g1064190	290	659
153	LG:1101065.1:2000SEP08	g2767633	422	482
153	LG:1101065.1:2000SEP08	g4196263	413	852
153	LG:1101065.1:2000SEP08	5622355H1	454	687
153	LG:1101065.1:2000SEP08	1275595F6	565	762
153	LG:1101065.1:2000SEP08	2626091H1	533	668
153	LG:1101065.1:2000SEP08	6339063H1	524	762
153	LG:1101065.1:2000SEP08	5907243H1	523	706
153	LG:1101065.1:2000SEP08	4254095H1	529	801
153	LG:1101065.1:2000SEP08	2417130H1	473	560
153	LG:1101065.1:2000SEP08	5095085H1	511	651
153	LG:1101065.1:2000SEP08	6413252H1	439	546
153	LG:1101065.1:2000SEP08	1298619H1	444	617
153	LG:1101065.1:2000SEP08	7030006H1	482	661
153	LG:1101065.1:2000SEP08	g678593	468	599
153	LG:1101065.1:2000SEP08	g3754064	502	924
153	LG:1101065.1:2000SEP08	6487688H1	224	751
153	LG:1101065.1:2000SEP08	6867273H1	225	738
153	LG:1101065.1:2000SEP08	g1977679	218	539
153	LG:1101065.1:2000SEP08	144037H1	237	547
153	LG:1101065.1:2000SEP08	g2397334	190	302
153	LG:1101065.1:2000SEP08	g869305	179	497
153	LG:1101065.1:2000SEP08	g274251	242	547
153	LG:1101065.1:2000SEP08	g680053	338	626
153	LG:1101065.1:2000SEP08	6927538H1	303	627
153	LG:1101065.1:2000SEP08	5045437H1	333	587
153	LG:1101065.1:2000SEP08	2443970H1	325	480
153	LG:1101065.1:2000SEP08	2443970F6	330	480
153	LG:1101065.1:2000SEP08	7092791H1	337	726
153	LG:1101065.1:2000SEP08	g4457572	352	617
153	LG:1101065.1:2000SEP08	4911166H1	366	614

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
153	LG:1101065.1:2000SEP08	g2180174	593	762
153	LG:1101065.1:2000SEP08	g2818383	605	729
153	LG:1101065.1:2000SEP08	5539050H1	604	740
153	LG:1101065.1:2000SEP08	4829460H1	647	894
153	LG:1101065.1:2000SEP08	3236041H1	572	663
153	LG:1101065.1:2000SEP08	3684012T6	566	667
153	LG:1101065.1:2000SEP08	7639253H1	116	546
153	LG:1101065.1:2000SEP08	2603820H1	114	306
153	LG:1101065.1:2000SEP08	g5673583	111	198
153	LG:1101065.1:2000SEP08	2603820F6	124	265
153	LG:1101065.1:2000SEP08	6583301T1	146	464
153	LG:1101065.1:2000SEP08	g824040	179	512
153	LG:1101065.1:2000SEP08	6908448J1	1	606
153	LG:1101065.1:2000SEP08	1275595H1	565	801
153	LG:1101065.1:2000SEP08	g1395972	562	664
153	LG:1101065.1:2000SEP08	5080092H1	562	665
153	LG:1101065.1:2000SEP08	2421749H1	566	813
153	LG:1101065.1:2000SEP08	g994630	566	845
153	LG:1101065.1:2000SEP08	6044326J1	407	664
153	LG:1101065.1:2000SEP08	7383747H1	380	631
153	LG:1101065.1:2000SEP08	6753293H1	370	542
153	LG:1101065.1:2000SEP08	g3144303	394	505
153	LG:1101065.1:2000SEP08	g1577192	398	577
153	LG:1101065.1:2000SEP08	g2539844	397	640
153	LG:1101065.1:2000SEP08	g2986729	434	923
153	LG:1101065.1:2000SEP08	4949289H1	393	665
153	LG:1101065.1:2000SEP08	1454401H1	418	599
153	LG:1101065.1:2000SEP08	g2629619	399	575
154	LG:475629.1:2000SEP08	5912746T9	14	469
154	LG:475629.1:2000SEP08	5912746H1	1	289
154	LG:475629.1:2000SEP08	5912746F6	1	570
154	LG:475629.1:2000SEP08	5912746F8	6	585
155	LI:348991.1:2000SEP08	7187352H1	1	154
155	LI:348991.1:2000SEP08	2156010H1	1	269
155	LI:348991.1:2000SEP08	2156010F6	1	437
155	LI:348991.1:2000SEP08	2157261F6	44	388
155	LI:348991.1:2000SEP08	2157261H1	44	295
155	LI:348991.1:2000SEP08	3494358H1	101	385
155	LI:348991.1:2000SEP08	1943827H1	131	375
155	LI:348991.1:2000SEP08	6296709H1	194	555
155	LI:348991.1:2000SEP08	2157261T6	373	906
155	LI:348991.1:2000SEP08	2156010T6	404	911
155	LI:348991.1:2000SEP08	g2158712	449	949
155	LI:348991.1:2000SEP08	g2963909	489	949
155	LI:348991.1:2000SEP08	g3752211	497	944
155	LI:348991.1:2000SEP08	7279628H1	511	822
155	LI:348991.1:2000SEP08	4246908H1	618	846
155	LI:348991.1:2000SEP08	1266194H1	713	938
155	LI:348991.1:2000SEP08	1266194F6	713	938

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
155	LI:348991.1:2000SEP08	1266194T6	725	898
156	LI:475629.1:2000SEP08	4408617H1	1	276
156	LI:475629.1:2000SEP08	5912746F8	6	588
156	LI:475629.1:2000SEP08	5912746T9	14	472
156	LI:475629.1:2000SEP08	5912746H1	1	290
156	LI:475629.1:2000SEP08	5912746F6	1	573
157	LI:261331.1:2000SEP08	4801843F8	1	451
157	LI:261331.1:2000SEP08	4802243H1	1	79
157	LI:261331.1:2000SEP08	4801843T8	50	259
158	LI:815686.1:2000SEP08	2750483R6	1559	1839
158	LI:815686.1:2000SEP08	g6040393	1593	1964
158	LI:815686.1:2000SEP08	2750483H1	1638	1839
158	LI:815686.1:2000SEP08	71679116V1	115	340
158	LI:815686.1:2000SEP08	3360770F6	316	760
158	LI:815686.1:2000SEP08	3360062T6	345	784
158	LI:815686.1:2000SEP08	3360778H1	491	760
158	LI:815686.1:2000SEP08	7766821H1	1522	2059
158	LI:815686.1:2000SEP08	g1993084	1559	1773
158	LI:815686.1:2000SEP08	7195403H1	633	1120
158	LI:815686.1:2000SEP08	3354351H1	689	802
158	LI:815686.1:2000SEP08	3234610H1	1271	1512
158	LI:815686.1:2000SEP08	7335384H1	1223	1673
158	LI:815686.1:2000SEP08	g3173294	1346	1658
158	LI:815686.1:2000SEP08	3512588H1	1351	1585
158	LI:815686.1:2000SEP08	4871079H1	1357	1614
158	LI:815686.1:2000SEP08	1255919H1	1401	1628
158	LI:815686.1:2000SEP08	7630779J1	1108	1658
158	LI:815686.1:2000SEP08	5771975H1	1156	1658
158	LI:815686.1:2000SEP08	7323904H2	1166	1658
158	LI:815686.1:2000SEP08	7335191H1	1223	1658
158	LI:815686.1:2000SEP08	2837075H1	1	90
158	LI:815686.1:2000SEP08	g4085228	1360	1658
158	LI:815686.1:2000SEP08	2837075F6	1	452
158	LI:815686.1:2000SEP08	5497582H1	1093	1259
158	LI:815686.1:2000SEP08	5717883H1	691	1160
158	LI:815686.1:2000SEP08	2822529H1	763	980
158	LI:815686.1:2000SEP08	7384792H1	833	1382
158	LI:815686.1:2000SEP08	2659590F6	996	1495
158	LI:815686.1:2000SEP08	2659590H1	996	1249
158	LI:815686.1:2000SEP08	5619301H1	1027	1306
158	LI:815686.1:2000SEP08	5497582F6	1093	1479
159	LI:1167327.2:2000SEP08	587588T6	147	665
159	LI:1167327.2:2000SEP08	5277673F8	162	318
159	LI:1167327.2:2000SEP08	6311748F8	238	318
159	LI:1167327.2:2000SEP08	5277673T8	162	318
159	LI:1167327.2:2000SEP08	2839106T6	89	677
159	LI:1167327.2:2000SEP08	7660359J1	1	90
159	LI:1167327.2:2000SEP08	2807456F6	1	515
159	LI:1167327.2:2000SEP08	2807456H1	1	252

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
159	LI:1167327.2:2000SEP08	7630459J1	3	483
159	LI:1167327.2:2000SEP08	587588H1	147	318
159	LI:1167327.2:2000SEP08	4625011T8	44	637
159	LI:1167327.2:2000SEP08	587588R6	147	489
159	LI:1167327.2:2000SEP08	2406278H1	400	630
159	LI:1167327.2:2000SEP08	2807456T6	122	678
159	LI:1167327.2:2000SEP08	8051364J1	140	647
160	LI:758009.3:2000SEP08	3903835R9	1	350
160	LI:758009.3:2000SEP08	3903772H1	1	281
160	LI:758009.3:2000SEP08	3903835H1	1	278
160	LI:758009.3:2000SEP08	3903772R9	2	574
160	LI:758009.3:2000SEP08	3903772R8	8	101
160	LI:758009.3:2000SEP08	3901518H1	257	573
160	LI:758009.3:2000SEP08	3902718H1	257	537
160	LI:758009.3:2000SEP08	3899759R8	257	860
160	LI:758009.3:2000SEP08	2431406H1	744	932
161	LG:331593.1:2000SEP08	336055T6	801	997
161	LG:331593.1:2000SEP08	7349784H1	855	1045
161	LG:331593.1:2000SEP08	g4310473	884	1045
161	LG:331593.1:2000SEP08	1840186F6	1	505
161	LG:331593.1:2000SEP08	1840186H1	10	261
161	LG:331593.1:2000SEP08	2435755H1	22	277
161	LG:331593.1:2000SEP08	656258H1	38	305
161	LG:331593.1:2000SEP08	1391084H1	50	272
161	LG:331593.1:2000SEP08	2433228H1	95	342
161	LG:331593.1:2000SEP08	1838010H1	188	467
161	LG:331593.1:2000SEP08	2436461H1	241	466
161	LG:331593.1:2000SEP08	1626212H1	374	580
161	LG:331593.1:2000SEP08	1626212F6	374	825
161	LG:331593.1:2000SEP08	408353H1	398	627
161	LG:331593.1:2000SEP08	334529H1	409	661
161	LG:331593.1:2000SEP08	1626212T6	442	718
161	LG:331593.1:2000SEP08	527671H1	609	854
161	LG:331593.1:2000SEP08	g5879650	620	1045
161	LG:331593.1:2000SEP08	g4390729	631	1051
161	LG:331593.1:2000SEP08	591704H1	689	914
161	LG:331593.1:2000SEP08	g4111182	719	1045
161	LG:331593.1:2000SEP08	g3330192	732	1045
161	LG:331593.1:2000SEP08	g3401889	735	1045
161	LG:331593.1:2000SEP08	336055R6	801	1045
161	LG:331593.1:2000SEP08	336055H1	801	1052
162	LI:1094174.1:2000SEP08	034344H1	578	759
162	LI:1094174.1:2000SEP08	3386726H1	605	903
162	LI:1094174.1:2000SEP08	3043448H1	296	547
162	LI:1094174.1:2000SEP08	3550695H1	297	600
162	LI:1094174.1:2000SEP08	g1373428	312	966
162	LI:1094174.1:2000SEP08	g1256065	313	954
162	LI:1094174.1:2000SEP08	3506620H1	226	545
162	LI:1094174.1:2000SEP08	3453506H1	226	547

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
162	LI:1094174.1:2000SEP08	3554818H1	226	540
162	LI:1094174.1:2000SEP08	319562H1	1608	1901
162	LI:1094174.1:2000SEP08	3726254H1	245	575
162	LI:1094174.1:2000SEP08	g1736042	1608	1903
162	LI:1094174.1:2000SEP08	5293714F9	240	655
162	LI:1094174.1:2000SEP08	3683647H1	678	993
162	LI:1094174.1:2000SEP08	70636185V1	670	1335
162	LI:1094174.1:2000SEP08	3536081H1	671	965
162	LI:1094174.1:2000SEP08	7212209H1	666	720
162	LI:1094174.1:2000SEP08	7206088H1	725	946
162	LI:1094174.1:2000SEP08	70528675V1	750	938
162	LI:1094174.1:2000SEP08	4541992H1	750	831
162	LI:1094174.1:2000SEP08	5182961H1	759	1015
162	LI:1094174.1:2000SEP08	5182934H1	759	1010
162	LI:1094174.1:2000SEP08	4642522H1	771	849
162	LI:1094174.1:2000SEP08	4411104H1	838	1129
162	LI:1094174.1:2000SEP08	4411111H1	839	1128
162	LI:1094174.1:2000SEP08	3569960H1	840	968
162	LI:1094174.1:2000SEP08	1219375H1	855	933
162	LI:1094174.1:2000SEP08	6164983H1	866	986
162	LI:1094174.1:2000SEP08	4073795H1	934	1226
162	LI:1094174.1:2000SEP08	4028074F6	935	1226
162	LI:1094174.1:2000SEP08	71302002V1	952	1563
162	LI:1094174.1:2000SEP08	71301519V1	1026	1572
162	LI:1094174.1:2000SEP08	1873863F6	1069	1429
162	LI:1094174.1:2000SEP08	1873863H1	1069	1342
162	LI:1094174.1:2000SEP08	4046595H1	1172	1451
162	LI:1094174.1:2000SEP08	71264717V1	1159	1400
162	LI:1094174.1:2000SEP08	598001R1	1199	1542
162	LI:1094174.1:2000SEP08	1539863H1	1251	1467
162	LI:1094174.1:2000SEP08	1539863R6	1252	1553
162	LI:1094174.1:2000SEP08	633495H1	1361	1565
162	LI:1094174.1:2000SEP08	1545185R6	1252	1556
162	LI:1094174.1:2000SEP08	6893605J1	1269	1772
162	LI:1094174.1:2000SEP08	70645278V1	1286	1999
162	LI:1094174.1:2000SEP08	4690701F6	1285	1693
162	LI:1094174.1:2000SEP08	5053162H1	1242	1365
162	LI:1094174.1:2000SEP08	5445131T9	1285	1703
162	LI:1094174.1:2000SEP08	3602270T6	1390	1949
162	LI:1094174.1:2000SEP08	345872H1	1307	1534
162	LI:1094174.1:2000SEP08	4549073T1	1313	1894
162	LI:1094174.1:2000SEP08	1881455H1	1356	1643
162	LI:1094174.1:2000SEP08	3646479H1	1287	1515
162	LI:1094174.1:2000SEP08	3114888H1	233	529
162	LI:1094174.1:2000SEP08	4549658H1	190	470
162	LI:1094174.1:2000SEP08	3554030H1	226	506
162	LI:1094174.1:2000SEP08	5334567H1	230	498
162	LI:1094174.1:2000SEP08	388101H1	1463	1721
162	LI:1094174.1:2000SEP08	71042310V1	1603	2001

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
162	LI:1094174.1:2000SEP08	4376379H1	1650	1901
162	LI:1094174.1:2000SEP08	3398162H1	108	338
162	LI:1094174.1:2000SEP08	2210409H1	226	523
162	LI:1094174.1:2000SEP08	877139H1	526	779
162	LI:1094174.1:2000SEP08	2237314H1	552	808
162	LI:1094174.1:2000SEP08	6099124H1	1513	1786
162	LI:1094174.1:2000SEP08	6099172H1	1513	1783
162	LI:1094174.1:2000SEP08	811786H1	1488	1563
162	LI:1094174.1:2000SEP08	375481H1	1515	1741
162	LI:1094174.1:2000SEP08	g2590198	1589	1900
162	LI:1094174.1:2000SEP08	2309787H1	109	363
162	LI:1094174.1:2000SEP08	2461733H1	525	751
162	LI:1094174.1:2000SEP08	4979849H1	232	503
162	LI:1094174.1:2000SEP08	3510749H1	226	564
162	LI:1094174.1:2000SEP08	1217823H1	1468	1715
162	LI:1094174.1:2000SEP08	4369748H1	1402	1646
162	LI:1094174.1:2000SEP08	2720340H1	105	366
162	LI:1094174.1:2000SEP08	4565946H1	551	702
162	LI:1094174.1:2000SEP08	512460H1	519	660
162	LI:1094174.1:2000SEP08	4348492H1	595	845
162	LI:1094174.1:2000SEP08	2898082H1	595	787
162	LI:1094174.1:2000SEP08	1986865H1	572	745
162	LI:1094174.1:2000SEP08	70645149V1	597	883
162	LI:1094174.1:2000SEP08	3052015H1	226	561
162	LI:1094174.1:2000SEP08	3551253H1	220	444
162	LI:1094174.1:2000SEP08	70645534V1	1	686
162	LI:1094174.1:2000SEP08	70634527V1	1	351
162	LI:1094174.1:2000SEP08	6830458H1	1	527
162	LI:1094174.1:2000SEP08	3166238H1	5	190
162	LI:1094174.1:2000SEP08	3616517H1	18	333
162	LI:1094174.1:2000SEP08	3616533H1	18	333
162	LI:1094174.1:2000SEP08	70636720V1	53	701
162	LI:1094174.1:2000SEP08	3042409H1	81	371
162	LI:1094174.1:2000SEP08	3042409F6	44	266
162	LI:1094174.1:2000SEP08	3051678H1	85	341
162	LI:1094174.1:2000SEP08	7255431H1	1560	1627
162	LI:1094174.1:2000SEP08	532131H1	237	505
162	LI:1094174.1:2000SEP08	868787H1	114	336
162	LI:1094174.1:2000SEP08	6118020H1	551	848
162	LI:1094174.1:2000SEP08	3485304H1	117	433
162	LI:1094174.1:2000SEP08	691288R6	115	497
162	LI:1094174.1:2000SEP08	578414H1	116	380
162	LI:1094174.1:2000SEP08	6708546J1	1926	2008
162	LI:1094174.1:2000SEP08	3509109H1	226	555
162	LI:1094174.1:2000SEP08	3510320H1	226	552
162	LI:1094174.1:2000SEP08	3552902H1	226	555
162	LI:1094174.1:2000SEP08	3551735H1	226	548
162	LI:1094174.1:2000SEP08	3881719H1	606	915
162	LI:1094174.1:2000SEP08	3751056H1	616	932

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
162	LI:1094174.1:2000SEP08	70636771V1	629	1135
162	LI:1094174.1:2000SEP08	70713233V1	640	933
162	LI:1094174.1:2000SEP08	657274H1	508	613
162	LI:1094174.1:2000SEP08	1213166H1	551	670
162	LI:1094174.1:2000SEP08	2239912H1	549	693
162	LI:1094174.1:2000SEP08	3502260H1	125	442
162	LI:1094174.1:2000SEP08	527310H1	124	382
162	LI:1094174.1:2000SEP08	4849338F9	258	610
162	LI:1094174.1:2000SEP08	3553862H1	231	560
162	LI:1094174.1:2000SEP08	3550548H1	251	569
162	LI:1094174.1:2000SEP08	3282780H1	471	757
162	LI:1094174.1:2000SEP08	193178H1	478	660
162	LI:1094174.1:2000SEP08	3842916H1	523	828
162	LI:1094174.1:2000SEP08	3486185H1	63	240
162	LI:1094174.1:2000SEP08	5188833F8	265	667
162	LI:1094174.1:2000SEP08	869151H1	114	335
162	LI:1094174.1:2000SEP08	3165580H1	63	286
162	LI:1094174.1:2000SEP08	4265782H1	226	530
162	LI:1094174.1:2000SEP08	71039331V1	1388	2019
162	LI:1094174.1:2000SEP08	2786711H1	265	547
162	LI:1094174.1:2000SEP08	3551187H1	267	506
162	LI:1094174.1:2000SEP08	1960243H1	268	526
162	LI:1094174.1:2000SEP08	3040931H1	270	561
162	LI:1094174.1:2000SEP08	4574306H1	256	541
162	LI:1094174.1:2000SEP08	7756579J1	276	422
162	LI:1094174.1:2000SEP08	5566655H1	290	543
162	LI:1094174.1:2000SEP08	5062073F8	271	361
162	LI:1094174.1:2000SEP08	3046233H1	290	586
162	LI:1094174.1:2000SEP08	4979008H1	236	555
162	LI:1094174.1:2000SEP08	1930187H1	110	373
162	LI:1094174.1:2000SEP08	3330307H1	109	404
162	LI:1094174.1:2000SEP08	5082521H1	63	267
162	LI:1094174.1:2000SEP08	1929426H1	107	345
162	LI:1094174.1:2000SEP08	4817280H1	227	464
162	LI:1094174.1:2000SEP08	4549828H1	226	399
162	LI:1094174.1:2000SEP08	3211703H1	227	414
162	LI:1094174.1:2000SEP08	5065535F9	304	972
162	LI:1094174.1:2000SEP08	4239104H1	233	347
162	LI:1094174.1:2000SEP08	5062227H1	228	463
162	LI:1094174.1:2000SEP08	3395083H1	230	546
162	LI:1094174.1:2000SEP08	3391535H1	230	539
162	LI:1094174.1:2000SEP08	3212986H1	231	539
162	LI:1094174.1:2000SEP08	4553293H1	234	321
162	LI:1094174.1:2000SEP08	1446532T6	1483	1951
162	LI:1094174.1:2000SEP08	3553810H1	226	559
162	LI:1094174.1:2000SEP08	70645812V1	153	720
162	LI:1094174.1:2000SEP08	5468431H1	226	398
162	LI:1094174.1:2000SEP08	70708072V1	1515	2000
162	LI:1094174.1:2000SEP08	1695993H1	107	320

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
162	LI:1094174.1:2000SEP08	g2288054	1711	1900
162	LI:1094174.1:2000SEP08	515388H1	1715	1915
162	LI:1094174.1:2000SEP08	6708673J1	1719	2001
162	LI:1094174.1:2000SEP08	g32232	1765	2001
162	LI:1094174.1:2000SEP08	502884H1	1751	1916
162	LI:1094174.1:2000SEP08	5526285H2	1800	2019
162	LI:1094174.1:2000SEP08	470702F1	1801	2000
162	LI:1094174.1:2000SEP08	g6038845	1835	2000
162	LI:1094174.1:2000SEP08	g5325998	1842	2000
162	LI:1094174.1:2000SEP08	g4892156	1851	2000
162	LI:1094174.1:2000SEP08	g6463789	1856	2000
162	LI:1094174.1:2000SEP08	g6464134	1862	2000
162	LI:1094174.1:2000SEP08	3077172H1	106	402
162	LI:1094174.1:2000SEP08	4142422H1	228	549
162	LI:1094174.1:2000SEP08	3396479H1	226	509
162	LI:1094174.1:2000SEP08	5334823H1	233	500
162	LI:1094174.1:2000SEP08	7374002H1	227	523
162	LI:1094174.1:2000SEP08	4882128T6	1487	1958
162	LI:1094174.1:2000SEP08	272691H1	1565	1779
162	LI:1094174.1:2000SEP08	7083449H1	332	807
162	LI:1094174.1:2000SEP08	2865092H1	261	588
162	LI:1094174.1:2000SEP08	2668233H1	258	528
162	LI:1094174.1:2000SEP08	3184648H1	226	549
162	LI:1094174.1:2000SEP08	3394068H1	342	641
162	LI:1094174.1:2000SEP08	1360670H1	368	643
162	LI:1094174.1:2000SEP08	70397513D1	381	692
162	LI:1094174.1:2000SEP08	4550009H1	391	647
162	LI:1094174.1:2000SEP08	4817019H1	392	662
162	LI:1094174.1:2000SEP08	7167291H1	410	505
162	LI:1094174.1:2000SEP08	70635270V1	434	1161
162	LI:1094174.1:2000SEP08	1711010H1	405	652
162	LI:1094174.1:2000SEP08	70635529V1	445	901
162	LI:1094174.1:2000SEP08	4977792H1	458	722
162	LI:1094174.1:2000SEP08	2259455H1	433	692
162	LI:1094174.1:2000SEP08	3388762H1	411	714
162	LI:1094174.1:2000SEP08	606770H1	435	692
162	LI:1094174.1:2000SEP08	2618137H1	437	685
162	LI:1094174.1:2000SEP08	4633948H1	437	708
162	LI:1094174.1:2000SEP08	746302H1	463	668
162	LI:1094174.1:2000SEP08	758784H1	472	703
162	LI:1094174.1:2000SEP08	g2005353	360	656
162	LI:1094174.1:2000SEP08	899393H1	349	483
162	LI:1094174.1:2000SEP08	2865820H1	114	419
162	LI:1094174.1:2000SEP08	3170360H1	112	367
162	LI:1094174.1:2000SEP08	3589722H1	109	419
162	LI:1094174.1:2000SEP08	71041521V1	1382	2019
162	LI:1094174.1:2000SEP08	g4534361	1379	1553
162	LI:1094174.1:2000SEP08	865371R1	236	835
162	LI:1094174.1:2000SEP08	1978636H1	1586	1843



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
162	LI:1094174.1:2000SEP08	2022771H1	1587	1838
162	LI:1094174.1:2000SEP08	2122657H1	1588	1854
162	LI:1094174.1:2000SEP08	g5662122	1637	1905
162	LI:1094174.1:2000SEP08	2647717H1	107	359
162	LI:1094174.1:2000SEP08	4569413H1	576	850
162	LI:1094174.1:2000SEP08	654311H1	1664	1944
162	LI:1094174.1:2000SEP08	7745687J1	1646	2002
162	LI:1094174.1:2000SEP08	g6199918	1689	1900
162	LI:1094174.1:2000SEP08	4468709T9	1658	1801
162	LI:1094174.1:2000SEP08	1814501H1	1707	1952
162	LI:1094174.1:2000SEP08	3554630H1	226	538
162	LI:1094174.1:2000SEP08	3552318H1	254	542
162	LI:1094174.1:2000SEP08	1662025H1	188	387
162	LI:1094174.1:2000SEP08	1680616H1	200	426
162	LI:1094174.1:2000SEP08	4550766H1	344	509
162	LI:1094174.1:2000SEP08	g804983	385	952
162	LI:1094174.1:2000SEP08	1709736H1	340	599
162	LI:1094174.1:2000SEP08	3512049H1	341	630
162	LI:1094174.1:2000SEP08	71079879V1	411	1095
162	LI:1094174.1:2000SEP08	604369H1	354	617
162	LI:1094174.1:2000SEP08	4768677H1	351	641
162	LI:1094174.1:2000SEP08	5090022H1	359	637
162	LI:1094174.1:2000SEP08	5338391H1	226	358
162	LI:1094174.1:2000SEP08	5067004H1	227	531
162	LI:1094174.1:2000SEP08	2643116H1	114	373
162	LI:1094174.1:2000SEP08	3800884H1	112	427
162	LI:1094174.1:2000SEP08	3529253H1	112	394
162	LI:1094174.1:2000SEP08	4980496H1	234	545
162	LI:1094174.1:2000SEP08	4240977H1	561	896
162	LI:1094174.1:2000SEP08	4976810H1	242	544
162	LI:1094174.1:2000SEP08	3551150H1	235	578
162	LI:1094174.1:2000SEP08	3395871H1	214	531
162	LI:1094174.1:2000SEP08	4141082H1	227	527
162	LI:1094174.1:2000SEP08	2189281H1	227	515
162	LI:1094174.1:2000SEP08	3550461H1	227	385
162	LI:1094174.1:2000SEP08	3554678H1	228	561
162	LI:1094174.1:2000SEP08	3692413H1	228	555
162	LI:1094174.1:2000SEP08	3506875H1	228	539
162	LI:1094174.1:2000SEP08	3508195H1	228	551
162	LI:1094174.1:2000SEP08	4042550H1	228	443
162	LI:1094174.1:2000SEP08	71532365V1	303	954
162	LI:1094174.1:2000SEP08	1682935H1	228	364
162	LI:1094174.1:2000SEP08	4152519H1	238	348
162	LI:1094174.1:2000SEP08	4150818H1	230	527
162	LI:1094174.1:2000SEP08	868827H1	170	493
162	LI:1094174.1:2000SEP08	3553827H1	233	535
162	LI:1094174.1:2000SEP08	g2595060	1567	1848
162	LI:1094174.1:2000SEP08	2956993H1	1334	1461
162	LI:1094174.1:2000SEP08	4882128F6	1330	1764

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
162	LI:1094174.1:2000SEP08	71042537V1	1390	2001
162	LI:1094174.1:2000SEP08	2534911H2	231	535
162	LI:1094174.1:2000SEP08	3690850H1	232	440
162	LI:1094174.1:2000SEP08	5472786H1	242	541
162	LI:1094174.1:2000SEP08	7193944H1	223	854
162	LI:1094174.1:2000SEP08	2211879H1	235	486
162	LI:1094174.1:2000SEP08	4057662H1	243	540
162	LI:1094174.1:2000SEP08	3553794H1	225	537
162	LI:1094174.1:2000SEP08	4154224H1	245	537
162	LI:1094174.1:2000SEP08	142299H1	551	759
162	LI:1094174.1:2000SEP08	2859033T6	1405	1890
162	LI:1094174.1:2000SEP08	71537164V1	1487	2000
162	LI:1094174.1:2000SEP08	71544615V1	1480	2019
162	LI:1094174.1:2000SEP08	g2005080	1478	1767
162	LI:1094174.1:2000SEP08	2847276H1	1478	1740
162	LI:1094174.1:2000SEP08	4305934H1	1453	1739
162	LI:1094174.1:2000SEP08	71075719V1	1478	2001
162	LI:1094174.1:2000SEP08	4885087F6	1399	1915
162	LI:1094174.1:2000SEP08	2640869H1	106	347
162	LI:1094174.1:2000SEP08	3104936H1	221	536
162	LI:1094174.1:2000SEP08	3393912H1	221	509
162	LI:1094174.1:2000SEP08	4975516H1	224	547
162	LI:1094174.1:2000SEP08	3314421H1	223	504
162	LI:1094174.1:2000SEP08	5065479H1	219	469
162	LI:1094174.1:2000SEP08	4951542H2	224	548
162	LI:1094174.1:2000SEP08	4057368H1	225	320
162	LI:1094174.1:2000SEP08	4107650H1	224	297
162	LI:1094174.1:2000SEP08	1962214H1	220	534
162	LI:1094174.1:2000SEP08	3686005H1	220	418
162	LI:1094174.1:2000SEP08	2738904H1	192	442
162	LI:1094174.1:2000SEP08	7363325H1	225	829
162	LI:1094174.1:2000SEP08	4975658H1	219	505
162	LI:1094174.1:2000SEP08	3392927H1	218	492
162	LI:1094174.1:2000SEP08	71470362V1	226	926
163	LI:814362.1:2000SEP08	3395087H1	3	289
163	LI:814362.1:2000SEP08	623520H1	7	277
163	LI:814362.1:2000SEP08	4547352H1	10	304
163	LI:814362.1:2000SEP08	1435555H1	10	276
163	LI:814362.1:2000SEP08	4672922H1	10	290
163	LI:814362.1:2000SEP08	1804332H1	13	273
163	LI:814362.1:2000SEP08	4601705H1	16	309
163	LI:814362.1:2000SEP08	908789H1	14	314
163	LI:814362.1:2000SEP08	4611548H1	14	310
163	LI:814362.1:2000SEP08	2923658H1	15	315
163	LI:814362.1:2000SEP08	700178H1	16	308
163	LI:814362.1:2000SEP08	611367H1	16	292
163	LI:814362.1:2000SEP08	6099855H1	16	289
163	LI:814362.1:2000SEP08	711210H1	17	303
163	LI:814362.1:2000SEP08	1598486H1	19	223

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
163	LI:814362.1:2000SEP08	1819534H1	3	227
163	LI:814362.1:2000SEP08	1598492H1	19	223
163	LI:814362.1:2000SEP08	3551250H1	29	228
163	LI:814362.1:2000SEP08	2716885H1	1	275
163	LI:814362.1:2000SEP08	5697720H1	2	299
163	LI:814362.1:2000SEP08	6028666H1	10	199
163	LI:814362.1:2000SEP08	2681413H1	12	314
163	LI:814362.1:2000SEP08	1521845H1	16	217
163	LI:814362.1:2000SEP08	6574674H1	18	530
163	LI:814362.1:2000SEP08	7961019H1	109	742
163	LI:814362.1:2000SEP08	1665076H1	199	435
163	LI:814362.1:2000SEP08	4847789H1	24	273
163	LI:814362.1:2000SEP08	2877428H1	152	434
163	LI:814362.1:2000SEP08	1820316H1	10	257
163	LI:814362.1:2000SEP08	1363931H1	221	406
163	LI:814362.1:2000SEP08	1821419H1	10	247
163	LI:814362.1:2000SEP08	2681645H1	1	260
163	LI:814362.1:2000SEP08	5079161H1	345	577
163	LI:814362.1:2000SEP08	3095675H1	317	614
163	LI:814362.1:2000SEP08	103404H1	279	481
163	LI:814362.1:2000SEP08	1599707H1	387	577
163	LI:814362.1:2000SEP08	1667374H1	462	691
163	LI:814362.1:2000SEP08	1666574H1	462	678
163	LI:814362.1:2000SEP08	1657409T7	380	916
163	LI:814362.1:2000SEP08	5181367H2	1	94
163	LI:814362.1:2000SEP08	7711466J1	296	446
163	LI:814362.1:2000SEP08	71369710V1	1	500
163	LI:814362.1:2000SEP08	5182722H1	1	188
163	LI:814362.1:2000SEP08	4847449H2	2	271
163	LI:814362.1:2000SEP08	369817H1	1	288
163	LI:814362.1:2000SEP08	71368710V1	11	492
163	LI:814362.1:2000SEP08	704481H1	16	249
163	LI:814362.1:2000SEP08	4372448H1	16	224
163	LI:814362.1:2000SEP08	71366624V1	12	474
163	LI:814362.1:2000SEP08	2749676F6	12	463
163	LI:814362.1:2000SEP08	2749676H1	12	259
163	LI:814362.1:2000SEP08	2746469H1	13	246
163	LI:814362.1:2000SEP08	71364401V1	13	532
163	LI:814362.1:2000SEP08	71369452V1	22	429
163	LI:814362.1:2000SEP08	4849434H1	1	258
163	LI:814362.1:2000SEP08	697829H1	1	294
163	LI:814362.1:2000SEP08	2693520H1	1	263
163	LI:814362.1:2000SEP08	2884241H1	1	264
163	LI:814362.1:2000SEP08	1745631H1	1	259
163	LI:814362.1:2000SEP08	1598761H1	2	213
163	LI:814362.1:2000SEP08	2996639H1	2	271
163	LI:814362.1:2000SEP08	2922283H1	2	281
163	LI:814362.1:2000SEP08	6097790F8	22	492
164	LI:219542.1:2000SEP08	g6662609	217	543

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
164	LI:219542.1:2000SEP08	g3595119	109	545
164	LI:219542.1:2000SEP08	g3280361	105	545
164	LI:219542.1:2000SEP08	g3594967	86	543
164	LI:219542.1:2000SEP08	g4683518	281	542
164	LI:219542.1:2000SEP08	g3092294	455	542
164	LI:219542.1:2000SEP08	g7456004	179	542
164	LI:219542.1:2000SEP08	g4265877	89	541
164	LI:219542.1:2000SEP08	g6713694	150	541
164	LI:219542.1:2000SEP08	g3644340	166	540
164	LI:219542.1:2000SEP08	g2114760	160	540
164	LI:219542.1:2000SEP08	g7457239	94	540
164	LI:219542.1:2000SEP08	g3240846	281	540
164	LI:219542.1:2000SEP08	g3049325	167	540
164	LI:219542.1:2000SEP08	g2840498	287	540
164	LI:219542.1:2000SEP08	g3016971	356	540
164	LI:219542.1:2000SEP08	g2741531	112	537
164	LI:219542.1:2000SEP08	g4737180	280	537
164	LI:219542.1:2000SEP08	g3280132	364	537
164	LI:219542.1:2000SEP08	g3127456	171	537
164	LI:219542.1:2000SEP08	g3146546	321	510
164	LI:219542.1:2000SEP08	8032184J1	1	488
164	LI:219542.1:2000SEP08	g2115091	8	467
164	LI:219542.1:2000SEP08	4530181H1	20	276
165	LI:726197.1:2000SEP08	6271522T8	1	488
165	LI:726197.1:2000SEP08	6271522H2	1	448
165	LI:726197.1:2000SEP08	6271522F8	1	512
166	LI:1075314.1:2000SEP08	6796070H1	1	474
166	LI:1075314.1:2000SEP08	6791032H1	1	464
166	LI:1075314.1:2000SEP08	6791032F8	19	247
166	LI:1075314.1:2000SEP08	6794038H1	404	847
166	LI:1075314.1:2000SEP08	6791032T8	485	1044
167	LI:437883.1:2000SEP08	6792592F8	1	677
167	LI:437883.1:2000SEP08	6793015F8	1	670
167	LI:437883.1:2000SEP08	6793015H1	1	570
167	LI:437883.1:2000SEP08	6792592H1	1	385
167	LI:437883.1:2000SEP08	3895660H1	30	189
167	LI:437883.1:2000SEP08	6792592T8	155	732
167	LI:437883.1:2000SEP08	3858443H1	431	693
168	LG:336265.1:2000SEP08	g2270863	1186	1602
168	LG:336265.1:2000SEP08	839079R1	1216	1764
168	LG:336265.1:2000SEP08	839079H1	1216	1465
168	LG:336265.1:2000SEP08	1664219T6	1279	1757
168	LG:336265.1:2000SEP08	5530561H1	1322	1605
168	LG:336265.1:2000SEP08	2119851H1	1024	1272
168	LG:336265.1:2000SEP08	g2031655	1049	1409
168	LG:336265.1:2000SEP08	5103748H1	1071	1338
168	LG:336265.1:2000SEP08	7437891H1	1118	1593
168	LG:336265.1:2000SEP08	079773H1	1133	1368
168	LG:336265.1:2000SEP08	358159H1	1133	1363

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
168	LG:336265.1:2000SEP08	5696151T9	1138	1702
168	LG:336265.1:2000SEP08	1848266H1	1175	1447
168	LG:336265.1:2000SEP08	2119851T6	1180	1765
168	LG:336265.1:2000SEP08	5105238T6	1180	1770
168	LG:336265.1:2000SEP08	g2238490	1451	1813
168	LG:336265.1:2000SEP08	g3173396	1544	1802
168	LG:336265.1:2000SEP08	g5765951	1407	1808
168	LG:336265.1:2000SEP08	g5529794	1404	1804
168	LG:336265.1:2000SEP08	g3429629	1384	1803
168	LG:336265.1:2000SEP08	g3151977	1326	1802
168	LG:336265.1:2000SEP08	g5632106	1360	1809
168	LG:336265.1:2000SEP08	g4989371	1342	1802
168	LG:336265.1:2000SEP08	g3179690	1653	1803
168	LG:336265.1:2000SEP08	g4875680	1670	1803
168	LG:336265.1:2000SEP08	g2163532	1682	1803
168	LG:336265.1:2000SEP08	2119851R6	987	1481
168	LG:336265.1:2000SEP08	5743126H1	1018	1304
168	LG:336265.1:2000SEP08	5105238T9	1017	1697
168	LG:336265.1:2000SEP08	5105238F8	906	1377
168	LG:336265.1:2000SEP08	5105238F9	905	1397
168	LG:336265.1:2000SEP08	1947115H1	905	1100
168	LG:336265.1:2000SEP08	5105238F6	905	1354
168	LG:336265.1:2000SEP08	7746203H1	754	1257
168	LG:336265.1:2000SEP08	3554943H1	777	981
168	LG:336265.1:2000SEP08	g3958830	898	1343
168	LG:336265.1:2000SEP08	7252476H1	1	580
168	LG:336265.1:2000SEP08	g1949296	278	591
168	LG:336265.1:2000SEP08	7440025H1	322	499
168	LG:336265.1:2000SEP08	5696151H1	473	734
168	LG:336265.1:2000SEP08	5696151F9	473	1082
168	LG:336265.1:2000SEP08	1664219H1	629	785
168	LG:336265.1:2000SEP08	1664219F6	629	1056
169	LG:407788.2:2000SEP08	5767076H1	78	602
169	LG:407788.2:2000SEP08	7746161H1	478	1094
169	LG:407788.2:2000SEP08	6711212H1	888	1414
169	LG:407788.2:2000SEP08	g2165285	896	1319
169	LG:407788.2:2000SEP08	3355568F6	1	336
169	LG:407788.2:2000SEP08	3355568H1	1	275
170	LG:1326925.1:2000SEP08	6825695H1	1	469
170	LG:1326925.1:2000SEP08	6825695J1	95	675
170	LG:1326925.1:2000SEP08	5872244H1	457	747
170	LG:1326925.1:2000SEP08	3408804F6	528	814
170	LG:1326925.1:2000SEP08	3408804H1	528	799
170	LG:1326925.1:2000SEP08	3408804T6	533	829
170	LG:1326925.1:2000SEP08	g188871	725	983
170	LG:1326925.1:2000SEP08	6820923H1	865	1029
170	LG:1326925.1:2000SEP08	5522310R6	879	1143
170	LG:1326925.1:2000SEP08	7400516H1	1042	1430
171	LI:332655.2:2000SEP08	70096608V1	227	637

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
171	LI:332655.2:2000SEP08	70029991D1	243	727
171	LI:332655.2:2000SEP08	70094732V1	265	796
171	LI:332655.2:2000SEP08	3096881H1	265	401
171	LI:332655.2:2000SEP08	70032023D1	276	777
171	LI:332655.2:2000SEP08	3098886H1	275	541
171	LI:332655.2:2000SEP08	70030960D1	283	725
171	LI:332655.2:2000SEP08	70099139V1	312	844
171	LI:332655.2:2000SEP08	7600091H1	314	782
171	LI:332655.2:2000SEP08	70033443D1	323	796
171	LI:332655.2:2000SEP08	70032717D1	330	860
171	LI:332655.2:2000SEP08	70094884V1	349	750
171	LI:332655.2:2000SEP08	70097593V1	374	841
171	LI:332655.2:2000SEP08	70100180V1	390	802
171	LI:332655.2:2000SEP08	70031426D1	399	950
171	LI:332655.2:2000SEP08	70031765D1	400	958
171	LI:332655.2:2000SEP08	70032527D1	401	881
171	LI:332655.2:2000SEP08	70094406V1	433	854
171	LI:332655.2:2000SEP08	g1959326	479	914
171	LI:332655.2:2000SEP08	70030000D1	484	976
171	LI:332655.2:2000SEP08	5766510T8	484	616
171	LI:332655.2:2000SEP08	5070935H1	498	757
171	LI:332655.2:2000SEP08	70095137V1	506	919
171	LI:332655.2:2000SEP08	70094797V1	512	951
171	LI:332655.2:2000SEP08	70034268D1	540	1093
171	LI:332655.2:2000SEP08	70033906D1	568	1184
171	LI:332655.2:2000SEP08	70094604V1	589	1099
171	LI:332655.2:2000SEP08	70099025V1	588	1066
171	LI:332655.2:2000SEP08	570522H1	606	876
171	LI:332655.2:2000SEP08	70096090V1	633	1180
171	LI:332655.2:2000SEP08	70031980D1	636	1144
171	LI:332655.2:2000SEP08	3872248T6	663	1142
171	LI:332655.2:2000SEP08	70030235D1	675	1092
171	LI:332655.2:2000SEP08	70034276D1	700	1144
171	LI:332655.2:2000SEP08	2850393T6	700	1142
171	LI:332655.2:2000SEP08	70033742D1	709	1144
171	LI:332655.2:2000SEP08	g2835010	716	1069
171	LI:332655.2:2000SEP08	70100118V1	734	1144
171	LI:332655.2:2000SEP08	70099040V1	737	1157
171	LI:332655.2:2000SEP08	70094459V1	739	1144
171	LI:332655.2:2000SEP08	70095263V1	753	1144
171	LI:332655.2:2000SEP08	70097249V1	763	1102
171	LI:332655.2:2000SEP08	70097115V1	764	1181
171	LI:332655.2:2000SEP08	70096990V1	782	1144
171	LI:332655.2:2000SEP08	70098090V1	832	1144
171	LI:332655.2:2000SEP08	g4763947	847	1144
171	LI:332655.2:2000SEP08	70095213V1	901	1370
171	LI:332655.2:2000SEP08	70097499V1	1	333
171	LI:332655.2:2000SEP08	70096744V1	1	351
171	LI:332655.2:2000SEP08	70095985V1	1	353

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
171	LI:332655.2:2000SEP08	70099142V1	1	440
171	LI:332655.2:2000SEP08	2850393F6	1	499
171	LI:332655.2:2000SEP08	70099114V1	1	442
171	LI:332655.2:2000SEP08	2850393H1	2	289
171	LI:332655.2:2000SEP08	861289H1	3	241
171	LI:332655.2:2000SEP08	70097343V1	3	476
171	LI:332655.2:2000SEP08	70098803V1	3	563
171	LI:332655.2:2000SEP08	2287640R6	5	485
171	LI:332655.2:2000SEP08	2287640H1	5	255
171	LI:332655.2:2000SEP08	3872248H1	11	301
171	LI:332655.2:2000SEP08	70032638D1	11	547
171	LI:332655.2:2000SEP08	3872248F6	11	517
171	LI:332655.2:2000SEP08	70095217V1	19	467
171	LI:332655.2:2000SEP08	70096220V1	44	487
171	LI:332655.2:2000SEP08	70032006D1	67	512
171	LI:332655.2:2000SEP08	70096911V1	82	538
171	LI:332655.2:2000SEP08	70032045D1	107	611
171	LI:332655.2:2000SEP08	70031433D1	110	658
171	LI:332655.2:2000SEP08	3403385H1	114	366
171	LI:332655.2:2000SEP08	1821848H1	133	370
171	LI:332655.2:2000SEP08	3531915H1	153	445
171	LI:332655.2:2000SEP08	2059050H1	160	405
171	LI:332655.2:2000SEP08	5742858H1	164	474
171	LI:332655.2:2000SEP08	2733553H1	185	415
171	LI:332655.2:2000SEP08	5025820H1	188	470
171	LI:332655.2:2000SEP08	70099520V1	192	521
171	LI:332655.2:2000SEP08	g1947735	206	551
171	LI:332655.2:2000SEP08	70095482V1	907	1144
171	LI:332655.2:2000SEP08	70095836V1	915	1144
171	LI:332655.2:2000SEP08	70099681V1	928	1144
171	LI:332655.2:2000SEP08	g4986367	978	1144
171	LI:332655.2:2000SEP08	g5661482	980	1142
171	LI:332655.2:2000SEP08	516193H1	1027	1136
172	LI:1184621.4:2000SEP08	8014017H1	295	895
172	LI:1184621.4:2000SEP08	7746161H1	484	1101
172	LI:1184621.4:2000SEP08	3355568H1	1	280
172	LI:1184621.4:2000SEP08	5767076H1	78	608
172	LI:1184621.4:2000SEP08	3355568F6	1	342
173	LI:2051386.1:2000SEP08	g1988131	86	329
173	LI:2051386.1:2000SEP08	8038751H1	1	616
173	LI:2051386.1:2000SEP08	4992337H1	204	321
173	LI:2051386.1:2000SEP08	3433251H1	203	459
173	LI:2051386.1:2000SEP08	6893208J1	349	658
173	LI:2051386.1:2000SEP08	5960485H1	386	793
173	LI:2051386.1:2000SEP08	g1383226	514	950
173	LI:2051386.1:2000SEP08	7613311J1	624	1207
173	LI:2051386.1:2000SEP08	g1383397	827	1165
174	LG:362757.1:2000SEP08	4084603H1	1	242
174	LG:362757.1:2000SEP08	6273740F8	1	673

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
174	LG:362757.1:2000SEP08	6274096H1	1	473
174	LG:362757.1:2000SEP08	6274096T8	28	674
175	LG:406770.1:2000SEP08	g4983713	1990	2080
175	LG:406770.1:2000SEP08	7010884H1	1171	1471
175	LG:406770.1:2000SEP08	938514R6	1	258
175	LG:406770.1:2000SEP08	938514H1	1	320
175	LG:406770.1:2000SEP08	938514R1	1	498
175	LG:406770.1:2000SEP08	4176087H1	110	414
175	LG:406770.1:2000SEP08	g1058461	299	571
175	LG:406770.1:2000SEP08	g1058474	300	538
175	LG:406770.1:2000SEP08	g2278147	1750	2097
175	LG:406770.1:2000SEP08	g1687124	1945	2078
175	LG:406770.1:2000SEP08	g1687230	1969	2078
175	LG:406770.1:2000SEP08	g1194732	1806	2085
175	LG:406770.1:2000SEP08	g3721294	1871	2082
175	LG:406770.1:2000SEP08	630365H1	1898	2078
175	LG:406770.1:2000SEP08	g1058378	1091	1471
175	LG:406770.1:2000SEP08	g1229629	1109	1482
175	LG:406770.1:2000SEP08	4981119H1	1003	1266
175	LG:406770.1:2000SEP08	4981119F6	1003	1568
175	LG:406770.1:2000SEP08	g6025310	1012	1474
175	LG:406770.1:2000SEP08	g1058364	1090	1472
175	LG:406770.1:2000SEP08	235785R6	712	917
175	LG:406770.1:2000SEP08	235785H1	712	915
175	LG:406770.1:2000SEP08	584125H1	799	1085
175	LG:406770.1:2000SEP08	7216683H1	835	1366
175	LG:406770.1:2000SEP08	284730H1	843	1046
175	LG:406770.1:2000SEP08	g3148103	1110	1473
175	LG:406770.1:2000SEP08	1886580H1	1193	1473
175	LG:406770.1:2000SEP08	g1225171	1373	1473
175	LG:406770.1:2000SEP08	6534885H1	1412	1949
175	LG:406770.1:2000SEP08	4050307T6	1469	2042
175	LG:406770.1:2000SEP08	g1687281	1482	1975
175	LG:406770.1:2000SEP08	g1687155	1482	1975
175	LG:406770.1:2000SEP08	g4896587	1600	2061
175	LG:406770.1:2000SEP08	235785T6	1627	2044
175	LG:406770.1:2000SEP08	g2277685	1723	2083
175	LG:406770.1:2000SEP08	g1272076	613	984
175	LG:406770.1:2000SEP08	7610621H1	397	941
175	LG:406770.1:2000SEP08	4050307F6	416	892
175	LG:406770.1:2000SEP08	4050307H1	416	711
175	LG:406770.1:2000SEP08	541736H1	451	678
176	LG:1094640.1:2000SEP08	4632658F6	1	436
176	LG:1094640.1:2000SEP08	4632658T6	254	704
176	LG:1094640.1:2000SEP08	g1784417	467	723
176	LG:1094640.1:2000SEP08	3011594H1	547	804
176	LG:1094640.1:2000SEP08	3011594F6	547	814
177	LG:001929.1:2000SEP08	7602642J1	504	1104
177	LG:001929.1:2000SEP08	g6570839	604	1029



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
177	LG:001929.1:2000SEP08	3384076H1	797	1044
177	LG:001929.1:2000SEP08	3384076F6	797	1383
177	LG:001929.1:2000SEP08	3401709H1	942	1198
177	LG:001929.1:2000SEP08	5555135T9	1084	1669
177	LG:001929.1:2000SEP08	3384169H1	1140	1381
177	LG:001929.1:2000SEP08	3384076T6	1151	1744
177	LG:001929.1:2000SEP08	2026773H1	1198	1466
177	LG:001929.1:2000SEP08	g2834870	1397	1765
177	LG:001929.1:2000SEP08	g3665721	1445	1772
177	LG:001929.1:2000SEP08	3384548H1	12	257
177	LG:001929.1:2000SEP08	7602642H1	18	195
177	LG:001929.1:2000SEP08	7604412J1	325	813
177	LG:001929.1:2000SEP08	7604412H1	326	777
177	LG:001929.1:2000SEP08	5555135H1	425	592
177	LG:001929.1:2000SEP08	5555135F9	428	832
177	LG:001929.1:2000SEP08	6808777J1	496	782
177	LG:001929.1:2000SEP08	6808777H1	497	782
177	LG:001929.1:2000SEP08	6808777F8	498	782
177	LG:001929.1:2000SEP08	6808777R8	496	800
177	LG:001929.1:2000SEP08	3402892H1	1	234
177	LG:001929.1:2000SEP08	3402892F6	1	611
177	LG:001929.1:2000SEP08	3385139H1	2	244
177	LG:001929.1:2000SEP08	3403431H1	3	243
177	LG:001929.1:2000SEP08	3385139F6	4	551
177	LG:001929.1:2000SEP08	3098178H1	12	267
177	LG:001929.1:2000SEP08	3098178F6	12	506
177	LG:001929.1:2000SEP08	3384548F6	12	432
177	LG:001929.1:2000SEP08	192201H1	1618	1827
178	LI:401322.1:2000SEP08	750543H1	1	185
178	LI:401322.1:2000SEP08	g1474315	44	326
178	LI:401322.1:2000SEP08	3722596H1	61	336
178	LI:401322.1:2000SEP08	4875149F6	70	197
178	LI:401322.1:2000SEP08	70496203V1	70	449
178	LI:401322.1:2000SEP08	4875149H1	70	259
178	LI:401322.1:2000SEP08	6561179H1	71	635
178	LI:401322.1:2000SEP08	2029094H1	122	375
178	LI:401322.1:2000SEP08	5950410H1	127	404
178	LI:401322.1:2000SEP08	5950630H1	127	408
178	LI:401322.1:2000SEP08	70498482V1	130	403
178	LI:401322.1:2000SEP08	70498458V1	131	374
178	LI:401322.1:2000SEP08	g4998660	182	481
178	LI:401322.1:2000SEP08	70497653V1	225	505
178	LI:401322.1:2000SEP08	70503545V1	234	462
178	LI:401322.1:2000SEP08	4933727H1	241	387
178	LI:401322.1:2000SEP08	3533254H1	275	569
178	LI:401322.1:2000SEP08	70496056V1	311	815
178	LI:401322.1:2000SEP08	70617675V1	322	597
178	LI:401322.1:2000SEP08	4875149T6	347	759
178	LI:401322.1:2000SEP08	70642080V1	378	636

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
178	LI:401322.1:2000SEP08	70499609V1	394	532
178	LI:401322.1:2000SEP08	70495192V1	467	817
178	LI:401322.1:2000SEP08	g3202357	462	808
178	LI:401322.1:2000SEP08	748032H1	581	800
179	LI:208748.1:2000SEP08	g5849939	2047	2505
179	LI:208748.1:2000SEP08	g5886544	2052	2505
179	LI:208748.1:2000SEP08	g7316701	2056	2509
179	LI:208748.1:2000SEP08	5547767T8	2062	2494
179	LI:208748.1:2000SEP08	g3897094	2067	2506
179	LI:208748.1:2000SEP08	g5589572	2075	2507
179	LI:208748.1:2000SEP08	7363924H1	1	421
179	LI:208748.1:2000SEP08	7216253H1	16	372
179	LI:208748.1:2000SEP08	2997925H1	152	412
179	LI:208748.1:2000SEP08	7216779H1	182	673
179	LI:208748.1:2000SEP08	4029524H1	317	583
179	LI:208748.1:2000SEP08	7677411H1	394	749
179	LI:208748.1:2000SEP08	70880160V1	433	910
179	LI:208748.1:2000SEP08	6976066H1	615	1197
179	LI:208748.1:2000SEP08	70881073V1	616	1048
179	LI:208748.1:2000SEP08	5589484F6	679	1082
179	LI:208748.1:2000SEP08	5589484H1	679	878
179	LI:208748.1:2000SEP08	5547767F8	699	1119
179	LI:208748.1:2000SEP08	5547767H1	699	903
179	LI:208748.1:2000SEP08	7690958J1	711	1197
179	LI:208748.1:2000SEP08	70879145V1	870	1332
179	LI:208748.1:2000SEP08	70882772V1	1028	1600
179	LI:208748.1:2000SEP08	70879812V1	1037	1688
179	LI:208748.1:2000SEP08	2289869H1	1045	1287
179	LI:208748.1:2000SEP08	3572255F6	1066	1491
179	LI:208748.1:2000SEP08	3572255H1	1066	1372
179	LI:208748.1:2000SEP08	g572859	1099	1487
179	LI:208748.1:2000SEP08	2866482H1	1162	1459
179	LI:208748.1:2000SEP08	7409011H1	1201	1820
179	LI:208748.1:2000SEP08	70882160V1	1231	1870
179	LI:208748.1:2000SEP08	7207904H1	1303	1812
179	LI:208748.1:2000SEP08	4779792H1	1332	1597
179	LI:208748.1:2000SEP08	2314765R6	1351	1756
179	LI:208748.1:2000SEP08	2314765H1	1351	1588
179	LI:208748.1:2000SEP08	70882288V1	1361	1839
179	LI:208748.1:2000SEP08	1573702H1	1390	1615
179	LI:208748.1:2000SEP08	1217087H1	1398	1631
179	LI:208748.1:2000SEP08	2741444F6	1408	1746
179	LI:208748.1:2000SEP08	2741444H1	1408	1662
179	LI:208748.1:2000SEP08	70879122V1	1458	2010
179	LI:208748.1:2000SEP08	6123696H1	1464	1941
179	LI:208748.1:2000SEP08	6127196H1	1464	1926
179	LI:208748.1:2000SEP08	3448750H1	1465	1730
179	LI:208748.1:2000SEP08	4209522H1	1473	1759
179	LI:208748.1:2000SEP08	767485H1	1505	1746

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
179	LI:208748.1:2000SEP08	6288827H1	1530	1812
179	LI:208748.1:2000SEP08	7240771H1	1537	2100
179	LI:208748.1:2000SEP08	4406924H1	1541	1808
179	LI:208748.1:2000SEP08	3019766H1	1584	1862
179	LI:208748.1:2000SEP08	4942607H1	1618	1904
179	LI:208748.1:2000SEP08	2642519F6	1651	2088
179	LI:208748.1:2000SEP08	2642519H1	1651	1884
179	LI:208748.1:2000SEP08	252812H1	1671	1973
179	LI:208748.1:2000SEP08	70881747V1	1719	1901
179	LI:208748.1:2000SEP08	70882562V1	1741	2165
179	LI:208748.1:2000SEP08	5189967H1	1755	2027
179	LI:208748.1:2000SEP08	7933991H1	1782	2352
179	LI:208748.1:2000SEP08	70882068V1	1815	2267
179	LI:208748.1:2000SEP08	2314765T6	1846	2467
179	LI:208748.1:2000SEP08	999167H1	1852	2104
179	LI:208748.1:2000SEP08	4029524T8	1863	2407
179	LI:208748.1:2000SEP08	g6697993	1922	2505
179	LI:208748.1:2000SEP08	2794808F6	1931	2431
179	LI:208748.1:2000SEP08	2794808H1	1931	2193
179	LI:208748.1:2000SEP08	2741444T6	1939	2462
179	LI:208748.1:2000SEP08	g5913894	1947	2333
179	LI:208748.1:2000SEP08	70881171V1	1968	2507
179	LI:208748.1:2000SEP08	70882247V1	1987	2503
179	LI:208748.1:2000SEP08	g5396272	2002	2469
179	LI:208748.1:2000SEP08	g3674187	2005	2466
179	LI:208748.1:2000SEP08	6855636H1	2019	2522
179	LI:208748.1:2000SEP08	g5636152	2028	2507
179	LI:208748.1:2000SEP08	2794808T6	2032	2468
179	LI:208748.1:2000SEP08	g5636140	2035	2507
179	LI:208748.1:2000SEP08	3315324H2	2039	2284
179	LI:208748.1:2000SEP08	g2563964	2044	2507
179	LI:208748.1:2000SEP08	g5888147	2080	2505
179	LI:208748.1:2000SEP08	g3899761	2091	2507
179	LI:208748.1:2000SEP08	6912743J1	2094	2492
179	LI:208748.1:2000SEP08	g2458063	2103	2510
179	LI:208748.1:2000SEP08	g1140592	2103	2501
179	LI:208748.1:2000SEP08	g3932768	2171	2506
179	LI:208748.1:2000SEP08	g4739395	2179	2501
179	LI:208748.1:2000SEP08	544703H1	2189	2418
179	LI:208748.1:2000SEP08	g2805935	2196	2333
179	LI:208748.1:2000SEP08	g565556	2208	2506
179	LI:208748.1:2000SEP08	3572255T6	2275	2467
179	LI:208748.1:2000SEP08	g2432032	2342	2507
180	LI:407242.1:2000SEP08	1271760H1	1	237
180	LI:407242.1:2000SEP08	7292121H1	38	497
180	LI:407242.1:2000SEP08	3237762H1	104	309
180	LI:407242.1:2000SEP08	2862509H1	151	429
180	LI:407242.1:2000SEP08	2862509F6	151	580
180	LI:407242.1:2000SEP08	5457948H1	217	480

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
180	LI:407242.1:2000SEP08	065132H1	523	692
180	LI:407242.1:2000SEP08	6254643H1	532	776
180	LI:407242.1:2000SEP08	6987868H1	549	1091
180	LI:407242.1:2000SEP08	6984055H1	549	1016
180	LI:407242.1:2000SEP08	5089291H1	680	943
180	LI:407242.1:2000SEP08	3232664H2	765	1004
180	LI:407242.1:2000SEP08	g2001171	787	1056
180	LI:407242.1:2000SEP08	4285194H1	873	1154
180	LI:407242.1:2000SEP08	7180376H1	951	1496
180	LI:407242.1:2000SEP08	482811H1	979	1187
180	LI:407242.1:2000SEP08	483426H1	979	1208
180	LI:407242.1:2000SEP08	482811R6	979	1466
180	LI:407242.1:2000SEP08	g760448	1499	1738
180	LI:407242.1:2000SEP08	7158891H1	1631	2045
180	LI:407242.1:2000SEP08	5260775H1	1659	1901
180	LI:407242.1:2000SEP08	g4902429	1757	1897
180	LI:407242.1:2000SEP08	7206772H1	1880	2043
180	LI:407242.1:2000SEP08	894787H1	1914	2027
180	LI:407242.1:2000SEP08	2862509T6	1164	1533
180	LI:407242.1:2000SEP08	6771708H1	1198	1755
180	LI:407242.1:2000SEP08	g4525016	1223	1533
180	LI:407242.1:2000SEP08	g4395234	1229	1533
180	LI:407242.1:2000SEP08	4041841H1	1287	1561
180	LI:407242.1:2000SEP08	743874H1	1330	1577
180	LI:407242.1:2000SEP08	3355268H1	1342	1537
180	LI:407242.1:2000SEP08	2725709H1	1387	1635
180	LI:407242.1:2000SEP08	2725709F6	1387	1745
180	LI:407242.1:2000SEP08	1952718R6	1932	2027
180	LI:407242.1:2000SEP08	1952718H1	1932	2027
181	LI:403409.1:2000SEP08	71599134V1	772	1403
181	LI:403409.1:2000SEP08	70522973V1	793	1395
181	LI:403409.1:2000SEP08	71602119V1	891	1390
181	LI:403409.1:2000SEP08	6532643H1	803	1384
181	LI:403409.1:2000SEP08	70522446V1	733	1374
181	LI:403409.1:2000SEP08	70523188V1	696	1360
181	LI:403409.1:2000SEP08	70524769V1	653	1352
181	LI:403409.1:2000SEP08	71556912V1	924	1343
181	LI:403409.1:2000SEP08	70522614V1	765	1339
181	LI:403409.1:2000SEP08	70522594V1	717	1338
181	LI:403409.1:2000SEP08	6329688H1	705	1325
181	LI:403409.1:2000SEP08	71603353V1	768	1418
181	LI:403409.1:2000SEP08	71603455V1	1252	1415
181	LI:403409.1:2000SEP08	71601114V1	724	1408
181	LI:403409.1:2000SEP08	71603450V1	1284	1460
181	LI:403409.1:2000SEP08	71598741V1	892	1454
181	LI:403409.1:2000SEP08	70522814V1	774	1452
181	LI:403409.1:2000SEP08	70646182V1	896	1439
181	LI:403409.1:2000SEP08	71599649V1	913	1435
181	LI:403409.1:2000SEP08	71598104V1	822	1430

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
181	LI:403409.1:2000SEP08	70522541V1	755	1409
181	LI:403409.1:2000SEP08	70522436V1	771	1422
181	LI:403409.1:2000SEP08	71598878V1	781	1510
181	LI:403409.1:2000SEP08	71600060V1	923	1503
181	LI:403409.1:2000SEP08	70526615V1	947	1493
181	LI:403409.1:2000SEP08	71599305V1	1044	1491
181	LI:403409.1:2000SEP08	71603274V1	949	1492
181	LI:403409.1:2000SEP08	71601895V1	910	1488
181	LI:403409.1:2000SEP08	7068876H1	1069	1483
181	LI:403409.1:2000SEP08	70646150V1	1000	1479
181	LI:403409.1:2000SEP08	71602212V1	862	1475
181	LI:403409.1:2000SEP08	70522443V1	933	1622
181	LI:403409.1:2000SEP08	71603173V1	979	1606
181	LI:403409.1:2000SEP08	71598064V1	895	1581
181	LI:403409.1:2000SEP08	4068325T6	1018	1641
181	LI:403409.1:2000SEP08	71600461V1	918	1430
181	LI:403409.1:2000SEP08	71601649V1	849	1425
181	LI:403409.1:2000SEP08	71601959V1	1803	2288
181	LI:403409.1:2000SEP08	g1998847	1973	2275
181	LI:403409.1:2000SEP08	6869946H1	1717	2275
181	LI:403409.1:2000SEP08	g4510725	1831	2272
181	LI:403409.1:2000SEP08	3541104H1	1991	2272
181	LI:403409.1:2000SEP08	70525434V1	1668	2257
181	LI:403409.1:2000SEP08	71600071V1	1438	2211
181	LI:403409.1:2000SEP08	6298042H1	1871	2146
181	LI:403409.1:2000SEP08	70531485V1	1972	2140
181	LI:403409.1:2000SEP08	71570166V1	1946	2130
181	LI:403409.1:2000SEP08	71572989V1	1947	2130
181	LI:403409.1:2000SEP08	71599459V1	1419	2126
181	LI:403409.1:2000SEP08	7468129H1	1672	2123
181	LI:403409.1:2000SEP08	71599020V1	1377	2056
181	LI:403409.1:2000SEP08	71599030V1	1330	2056
181	LI:403409.1:2000SEP08	71600742V1	1513	2002
181	LI:403409.1:2000SEP08	70526344V1	1387	1988
181	LI:403409.1:2000SEP08	70522547V1	1396	1983
181	LI:403409.1:2000SEP08	71600018V1	1280	1965
181	LI:403409.1:2000SEP08	71598137V1	1757	1940
181	LI:403409.1:2000SEP08	6412487H1	1494	1931
181	LI:403409.1:2000SEP08	71559794V1	1463	1905
181	LI:403409.1:2000SEP08	71602804V1	1135	1832
181	LI:403409.1:2000SEP08	71600670V1	1402	1823
181	LI:403409.1:2000SEP08	5534143H1	1578	1820
181	LI:403409.1:2000SEP08	71602415V1	1308	1755
181	LI:403409.1:2000SEP08	71601115V1	1097	1726
181	LI:403409.1:2000SEP08	71600852V1	1086	1708
181	LI:403409.1:2000SEP08	70522029V1	1187	1709
181	LI:403409.1:2000SEP08	g1383466	1395	1696
181	LI:403409.1:2000SEP08	g3254782	1346	1682
181	LI:403409.1:2000SEP08	g1400213	1337	1682

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
181	LI:403409.1:2000SEP08	71597884V1	1113	1675
181	LI:403409.1:2000SEP08	70525805V1	1088	1643
181	LI:403409.1:2000SEP08	70522266V1	920	1644
181	LI:403409.1:2000SEP08	71602346V1	1075	1639
181	LI:403409.1:2000SEP08	70525328V1	1039	1641
181	LI:403409.1:2000SEP08	6141236T8	989	1579
181	LI:403409.1:2000SEP08	70523039V1	862	1577
181	LI:403409.1:2000SEP08	71599563V1	934	1577
181	LI:403409.1:2000SEP08	3256362H1	1308	1554
181	LI:403409.1:2000SEP08	71582870V1	1082	1322
181	LI:403409.1:2000SEP08	70523160V1	587	1316
181	LI:403409.1:2000SEP08	g6570650	881	1310
181	LI:403409.1:2000SEP08	71602205V1	830	1306
181	LI:403409.1:2000SEP08	71599405V1	861	1305
181	LI:403409.1:2000SEP08	71599960V1	738	1299
181	LI:403409.1:2000SEP08	71600217V1	738	1298
181	LI:403409.1:2000SEP08	71598742V1	553	1294
181	LI:403409.1:2000SEP08	70525194V1	724	1283
181	LI:403409.1:2000SEP08	71601307V1	716	1270
181	LI:403409.1:2000SEP08	70524986V1	585	1261
181	LI:403409.1:2000SEP08	70524908V1	682	1254
181	LI:403409.1:2000SEP08	70524071V1	682	1253
181	LI:403409.1:2000SEP08	71598545V1	680	1248
181	LI:403409.1:2000SEP08	6333709H1	705	1240
181	LI:403409.1:2000SEP08	7377730H1	697	1236
181	LI:403409.1:2000SEP08	71600305V1	716	1196
181	LI:403409.1:2000SEP08	71601960V1	627	1190
181	LI:403409.1:2000SEP08	71602858V1	513	1189
181	LI:403409.1:2000SEP08	71601208V1	616	1179
181	LI:403409.1:2000SEP08	7080559H1	760	1170
181	LI:403409.1:2000SEP08	71598389V1	650	1156
181	LI:403409.1:2000SEP08	71571578V1	747	1126
181	LI:403409.1:2000SEP08	71601452V1	535	1111
181	LI:403409.1:2000SEP08	71598584V1	402	1113
181	LI:403409.1:2000SEP08	4068325F6	515	1108
181	LI:403409.1:2000SEP08	71599561V1	412	1098
181	LI:403409.1:2000SEP08	71603370V1	588	1074
181	LI:403409.1:2000SEP08	71598865V1	713	1079
181	LI:403409.1:2000SEP08	71602389V1	520	1073
181	LI:403409.1:2000SEP08	71603466V1	374	1066
181	LI:403409.1:2000SEP08	71598103V1	595	1051
181	LI:403409.1:2000SEP08	71600023V1	485	1043
181	LI:403409.1:2000SEP08	71600063V1	486	1043
181	LI:403409.1:2000SEP08	71601379V1	339	1041
181	LI:403409.1:2000SEP08	71600010V1	353	1027
181	LI:403409.1:2000SEP08	71599332V1	386	1008
181	LI:403409.1:2000SEP08	71602290V1	485	982
181	LI:403409.1:2000SEP08	71599504V1	282	941
181	LI:403409.1:2000SEP08	70525699V1	339	890

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
181	LI:403409.1:2000SEP08	373638H1	663	880
181	LI:403409.1:2000SEP08	3751233H1	577	867
181	LI:403409.1:2000SEP08	5812190H1	644	845
181	LI:403409.1:2000SEP08	5812189H1	644	840
181	LI:403409.1:2000SEP08	3629455H1	560	835
181	LI:403409.1:2000SEP08	4068325H1	517	804
181	LI:403409.1:2000SEP08	311752H1	657	744
181	LI:403409.1:2000SEP08	70535025V1	354	705
181	LI:403409.1:2000SEP08	6141236F8	28	633
181	LI:403409.1:2000SEP08	70525881V1	102	628
181	LI:403409.1:2000SEP08	3409854H1	357	618
181	LI:403409.1:2000SEP08	6328822H1	47	592
181	LI:403409.1:2000SEP08	70522980V1	1	589
181	LI:403409.1:2000SEP08	71599955V1	1	538
181	LI:403409.1:2000SEP08	70524819V1	1	487
181	LI:403409.1:2000SEP08	3323733H1	192	461
181	LI:403409.1:2000SEP08	71601789V1	1	452
181	LI:403409.1:2000SEP08	71598908V1	1	435
181	LI:403409.1:2000SEP08	6134014H1	110	401
181	LI:403409.1:2000SEP08	6141236H1	28	370
181	LI:403409.1:2000SEP08	3388023H1	89	361
181	LI:403409.1:2000SEP08	2723940F6	1	333
181	LI:403409.1:2000SEP08	3126704H1	22	298
181	LI:403409.1:2000SEP08	g1998848	22	296
181	LI:403409.1:2000SEP08	5643224H1	22	275
181	LI:403409.1:2000SEP08	3256088H1	24	259
181	LI:403409.1:2000SEP08	5499626H1	59	250
181	LI:403409.1:2000SEP08	1728945H1	26	246
181	LI:403409.1:2000SEP08	2723940H1	1	243
181	LI:403409.1:2000SEP08	5500309H1	25	226
181	LI:403409.1:2000SEP08	5499909H1	1	174
181	LI:403409.1:2000SEP08	5500026H1	3	160
182	LI:450798.1:2000SEP08	70998432V1	348	871
182	LI:450798.1:2000SEP08	71300061V1	423	1033
182	LI:450798.1:2000SEP08	70996277V1	454	1126
182	LI:450798.1:2000SEP08	5906243F8	1	150
182	LI:450798.1:2000SEP08	71298861V1	1	381
182	LI:450798.1:2000SEP08	70997953V1	1252	1785
182	LI:450798.1:2000SEP08	70998492V1	1291	1779
182	LI:450798.1:2000SEP08	71299576V1	1276	1779
182	LI:450798.1:2000SEP08	70998290V1	1301	1754
182	LI:450798.1:2000SEP08	71001007V1	1487	1666
182	LI:450798.1:2000SEP08	71002412V1	1487	1665
182	LI:450798.1:2000SEP08	71284994V1	1622	1784
182	LI:450798.1:2000SEP08	70996506V1	1158	1779
182	LI:450798.1:2000SEP08	71299572V1	1134	1758
182	LI:450798.1:2000SEP08	70998139V1	1160	1808
182	LI:450798.1:2000SEP08	70996451V1	1143	1779
182	LI:450798.1:2000SEP08	71299204V1	694	1129

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
182	LI:450798.1:2000SEP08	70995811V1	707	1320
182	LI:450798.1:2000SEP08	70996755V1	711	1216
182	LI:450798.1:2000SEP08	71298736V1	719	1271
182	LI:450798.1:2000SEP08	71298890V1	735	1345
182	LI:450798.1:2000SEP08	70995249V1	749	1340
182	LI:450798.1:2000SEP08	70995662V1	1	528
182	LI:450798.1:2000SEP08	70998510V1	1	541
182	LI:450798.1:2000SEP08	70997263V1	269	814
182	LI:450798.1:2000SEP08	70995503V1	328	878
182	LI:450798.1:2000SEP08	70997710V1	344	979
182	LI:450798.1:2000SEP08	70996031V1	344	978
182	LI:450798.1:2000SEP08	70998404V1	1	491
182	LI:450798.1:2000SEP08	70995922V1	1	442
182	LI:450798.1:2000SEP08	5906243H1	1	293
182	LI:450798.1:2000SEP08	5906243F6	2	567
182	LI:450798.1:2000SEP08	70996045V1	194	666
182	LI:450798.1:2000SEP08	70997325V1	1144	1702
182	LI:450798.1:2000SEP08	70997726V1	1145	1791
182	LI:450798.1:2000SEP08	71004095V1	1143	1376
182	LI:450798.1:2000SEP08	71003787V1	1143	1374
182	LI:450798.1:2000SEP08	70998190V1	1173	1779
182	LI:450798.1:2000SEP08	70995151V1	755	960
182	LI:450798.1:2000SEP08	70997688V1	765	1497
182	LI:450798.1:2000SEP08	70997481V1	775	1374
182	LI:450798.1:2000SEP08	70997157V1	794	1372
182	LI:450798.1:2000SEP08	70954505V1	799	1033
182	LI:450798.1:2000SEP08	70954308V1	805	1204
182	LI:450798.1:2000SEP08	70995706V1	664	1316
182	LI:450798.1:2000SEP08	70996760V1	541	1112
182	LI:450798.1:2000SEP08	70996660V1	539	1114
182	LI:450798.1:2000SEP08	70997885V1	565	1098
182	LI:450798.1:2000SEP08	71003084V1	569	730
182	LI:450798.1:2000SEP08	6594478H1	647	1209
182	LI:450798.1:2000SEP08	70995050V1	651	1339
182	LI:450798.1:2000SEP08	71032609V1	495	1112
182	LI:450798.1:2000SEP08	70996627V1	807	1350
182	LI:450798.1:2000SEP08	70997576V1	831	1403
182	LI:450798.1:2000SEP08	70997754V1	846	1538
182	LI:450798.1:2000SEP08	71299090V1	850	1467
182	LI:450798.1:2000SEP08	70998737V1	850	1470
182	LI:450798.1:2000SEP08	70996977V1	935	1569
182	LI:450798.1:2000SEP08	71300385V1	938	1609
182	LI:450798.1:2000SEP08	71298713V1	972	1305
182	LI:450798.1:2000SEP08	71300270V1	985	1677
182	LI:450798.1:2000SEP08	5906243T9	1026	1666
182	LI:450798.1:2000SEP08	70995801V1	1077	1679
182	LI:450798.1:2000SEP08	70998862V1	1116	1779
182	LI:450798.1:2000SEP08	70998830V1	1156	1779
183	LI:410317.1:2000SEP08	4999885F6	264	643



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
183	LI:410317.1:2000SEP08	71740164V1	328	976
183	LI:410317.1:2000SEP08	71738350V1	599	1069
183	LI:410317.1:2000SEP08	156117H1	602	820
183	LI:410317.1:2000SEP08	4365223H1	637	907
183	LI:410317.1:2000SEP08	4365223F6	637	1165
183	LI:410317.1:2000SEP08	71737923V1	674	1333
183	LI:410317.1:2000SEP08	71739693V1	777	1355
183	LI:410317.1:2000SEP08	71736690V1	788	1321
183	LI:410317.1:2000SEP08	71740125V1	810	1393
183	LI:410317.1:2000SEP08	71735504V1	830	1497
183	LI:410317.1:2000SEP08	71736650V1	851	1499
183	LI:410317.1:2000SEP08	71735712V1	887	1128
183	LI:410317.1:2000SEP08	4365223T6	898	1414
183	LI:410317.1:2000SEP08	71739562V1	902	1515
183	LI:410317.1:2000SEP08	71735544V1	913	1500
183	LI:410317.1:2000SEP08	4361845H1	1226	1464
183	LI:410317.1:2000SEP08	4361845F6	1227	1457
183	LI:410317.1:2000SEP08	g3431989	1375	1463
183	LI:410317.1:2000SEP08	71735947V1	932	1529
183	LI:410317.1:2000SEP08	71740033V1	973	1496
183	LI:410317.1:2000SEP08	71738781V1	925	1395
183	LI:410317.1:2000SEP08	1805911T6	1043	1418
183	LI:410317.1:2000SEP08	4999571H1	1044	1328
183	LI:410317.1:2000SEP08	5139406H1	1143	1428
183	LI:410317.1:2000SEP08	1805911F6	1	222
183	LI:410317.1:2000SEP08	1805911H1	1	274
183	LI:410317.1:2000SEP08	1807624H1	1	260
183	LI:410317.1:2000SEP08	71738190V1	232	775
183	LI:410317.1:2000SEP08	71740974V1	232	768
183	LI:410317.1:2000SEP08	71735931V1	238	818
183	LI:410317.1:2000SEP08	71740141V1	238	825
183	LI:410317.1:2000SEP08	71739935V1	238	829
183	LI:410317.1:2000SEP08	71736466V1	238	768
183	LI:410317.1:2000SEP08	4995880H1	242	522
183	LI:410317.1:2000SEP08	4999885H1	238	408
184	LI:340268.1:2000SEP08	71072293V1	758	1308
184	LI:340268.1:2000SEP08	71069652V1	795	882
184	LI:340268.1:2000SEP08	71072824V1	813	882
184	LI:340268.1:2000SEP08	g4390671	1012	1289
184	LI:340268.1:2000SEP08	4532644H1	1	98
184	LI:340268.1:2000SEP08	4525543F6	1	207
184	LI:340268.1:2000SEP08	4525543H1	1	85
184	LI:340268.1:2000SEP08	5026361F7	110	632
184	LI:340268.1:2000SEP08	5026361H1	110	373
184	LI:340268.1:2000SEP08	71065144V1	152	237
184	LI:340268.1:2000SEP08	71071658V1	218	835
184	LI:340268.1:2000SEP08	71070531V1	321	559
184	LI:340268.1:2000SEP08	71263162V1	327	619
184	LI:340268.1:2000SEP08	71251415V1	429	905

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
184	LI:340268.1:2000SEP08	71071531V1	503	1191
184	LI:340268.1:2000SEP08	71069006V1	599	1223
184	LI:340268.1:2000SEP08	71252155V1	601	915
184	LI:340268.1:2000SEP08	71547677V1	666	1118
184	LI:340268.1:2000SEP08	71251537V1	687	1308
184	LI:340268.1:2000SEP08	4525543T6	716	1254
184	LI:340268.1:2000SEP08	6566268H1	723	920
184	LI:340268.1:2000SEP08	71251602V1	751	1306
185	LI:2051671.1:2000SEP08	71079934V1	231	892
185	LI:2051671.1:2000SEP08	71075860V1	279	848
185	LI:2051671.1:2000SEP08	71078989V1	357	921
185	LI:2051671.1:2000SEP08	g3665040	550	1038
185	LI:2051671.1:2000SEP08	g3134614	956	1038
185	LI:2051671.1:2000SEP08	2884229F6	830	1038
185	LI:2051671.1:2000SEP08	2884229T6	832	996
185	LI:2051671.1:2000SEP08	2884229H1	830	1038
185	LI:2051671.1:2000SEP08	g796618	685	1047
185	LI:2051671.1:2000SEP08	5609828H1	684	838
185	LI:2051671.1:2000SEP08	71078444V1	1	571
185	LI:2051671.1:2000SEP08	71080417V1	147	723
185	LI:2051671.1:2000SEP08	71080218V1	166	656
185	LI:2051671.1:2000SEP08	71078123V1	490	1056
185	LI:2051671.1:2000SEP08	71079055V1	494	1055
185	LI:2051671.1:2000SEP08	1852882H1	523	822
185	LI:2051671.1:2000SEP08	6426159H1	544	1055
185	LI:2051671.1:2000SEP08	g683142	809	1038
185	LI:2051671.1:2000SEP08	2883977F6	830	1038
185	LI:2051671.1:2000SEP08	3468787H1	951	1038
185	LI:2051671.1:2000SEP08	g5449833	687	1038
185	LI:2051671.1:2000SEP08	7953619H1	752	1056
185	LI:2051671.1:2000SEP08	71079993V1	475	1064
185	LI:2051671.1:2000SEP08	71134865V1	484	888
185	LI:2051671.1:2000SEP08	71077560V1	571	1038
185	LI:2051671.1:2000SEP08	600643H1	593	890
185	LI:2051671.1:2000SEP08	1702744F6	597	1038
185	LI:2051671.1:2000SEP08	g517763	563	1038
185	LI:2051671.1:2000SEP08	4048076H1	857	1010
185	LI:2051671.1:2000SEP08	1617868T6	551	996
185	LI:2051671.1:2000SEP08	71079428V1	569	1204
185	LI:2051671.1:2000SEP08	2202245T6	462	998
185	LI:2051671.1:2000SEP08	4415153H1	466	728
185	LI:2051671.1:2000SEP08	2199030T6	382	996
185	LI:2051671.1:2000SEP08	71129432V1	381	676
185	LI:2051671.1:2000SEP08	71078955V1	475	1066
185	LI:2051671.1:2000SEP08	1702744H1	597	822
185	LI:2051671.1:2000SEP08	71077372V1	168	673
185	LI:2051671.1:2000SEP08	1702744T6	598	995
185	LI:2051671.1:2000SEP08	70876374V1	213	493
185	LI:2051671.1:2000SEP08	71075993V1	210	853

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
186	LG:998844.1:2000SEP08	g6569256	1	394
186	LG:998844.1:2000SEP08	6764364J1	1	294
186	LG:998844.1:2000SEP08	g273443	2	422
187	LG:1043787.1:2000SEP08	6794461F8	1	591
187	LG:1043787.1:2000SEP08	6794461H1	2	449
187	LG:1043787.1:2000SEP08	6798630H1	24	445
187	LG:1043787.1:2000SEP08	6798630F8	24	481
187	LG:1043787.1:2000SEP08	6798660H1	24	340
187	LG:1043787.1:2000SEP08	6798660F8	24	591
187	LG:1043787.1:2000SEP08	6794461T8	79	732
187	LG:1043787.1:2000SEP08	6798660T8	241	696
188	LG:1098931.16:2000SEP08	7664285J1	1	582
188	LG:1098931.16:2000SEP08	1671805H1	291	515
189	LG:199423.2:2000SEP08	2757508H1	1	257
190	LI:1075297.1:2000SEP08	5334749F6	1	510
190	LI:1075297.1:2000SEP08	5334749F8	1	594
190	LI:1075297.1:2000SEP08	5334749H1	1	128
190	LI:1075297.1:2000SEP08	5334749T8	245	810
191	LI:1043321.1:2000SEP08	6796976H1	1	254
191	LI:1043321.1:2000SEP08	6796976F8	45	586
191	LI:1043321.1:2000SEP08	6796976T8	218	780
192	LI:297070.1:2000SEP08	7042156H1	457	984
192	LI:297070.1:2000SEP08	7766217J1	464	1125
192	LI:297070.1:2000SEP08	1564083H1	682	869
192	LI:297070.1:2000SEP08	g3676974	1	430
192	LI:297070.1:2000SEP08	2640925H1	228	476
192	LI:297070.1:2000SEP08	7766217H1	263	857
192	LI:297070.1:2000SEP08	7042156F8	457	983
192	LI:297070.1:2000SEP08	5539166F6	756	1096
192	LI:297070.1:2000SEP08	5539166H1	756	965
193	LI:1085041.1:2000SEP08	6796417H1	1	340
193	LI:1085041.1:2000SEP08	6796417F8	5	601
193	LI:1085041.1:2000SEP08	6796417T8	139	612
193	LI:1085041.1:2000SEP08	6793856T8	331	640
193	LI:1085041.1:2000SEP08	6793856H1	332	481
194	LI:1071544.1:2000SEP08	6792174F8	1	578
194	LI:1071544.1:2000SEP08	6792174H1	1	570
194	LI:1071544.1:2000SEP08	6792174T8	2	616
195	LI:2052480.1:2000SEP08	70844959V1	1025	1696
195	LI:2052480.1:2000SEP08	2243248H1	1114	1368
195	LI:2052480.1:2000SEP08	71045682V1	411	572
195	LI:2052480.1:2000SEP08	71223029V1	493	1123
195	LI:2052480.1:2000SEP08	g657732	539	917
195	LI:2052480.1:2000SEP08	70844094V1	376	1062
195	LI:2052480.1:2000SEP08	71224053V1	258	787
195	LI:2052480.1:2000SEP08	6787834H1	293	735
195	LI:2052480.1:2000SEP08	5072694H1	28	118
195	LI:2052480.1:2000SEP08	g3756288	1287	1719
195	LI:2052480.1:2000SEP08	4198326F6	1289	1816

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
195	LI:2052480.1:2000SEP08	2924556H1	895	1117
195	LI:2052480.1:2000SEP08	70846478V1	763	1171
195	LI:2052480.1:2000SEP08	2925554H1	1192	1483
195	LI:2052480.1:2000SEP08	g7045146	1249	1684
195	LI:2052480.1:2000SEP08	4198326H1	1289	1591
195	LI:2052480.1:2000SEP08	5072595H1	29	120
195	LI:2052480.1:2000SEP08	70846219V1	38	642
195	LI:2052480.1:2000SEP08	5072595F8	72	225
195	LI:2052480.1:2000SEP08	71223294V1	82	618
195	LI:2052480.1:2000SEP08	g4852852	792	950
195	LI:2052480.1:2000SEP08	70845403V1	817	1400
195	LI:2052480.1:2000SEP08	3696915F6	824	1268
195	LI:2052480.1:2000SEP08	3696915H1	824	1120
195	LI:2052480.1:2000SEP08	70845925V1	891	1417
195	LI:2052480.1:2000SEP08	70844782V1	740	1140
195	LI:2052480.1:2000SEP08	6934749H1	760	1346
195	LI:2052480.1:2000SEP08	71223146V1	553	1086
195	LI:2052480.1:2000SEP08	70844786V1	732	1359
195	LI:2052480.1:2000SEP08	g1721215	1326	1488
195	LI:2052480.1:2000SEP08	7424361T1	1323	1539
195	LI:2052480.1:2000SEP08	70845065V1	927	1522
195	LI:2052480.1:2000SEP08	71224205V1	936	1539
195	LI:2052480.1:2000SEP08	70842872V1	1019	1466
195	LI:2052480.1:2000SEP08	70845463V1	1291	1784
195	LI:2052480.1:2000SEP08	70845678V1	1147	1763
195	LI:2052480.1:2000SEP08	g657793	1	55
195	LI:2052480.1:2000SEP08	1501410F6	1	288
195	LI:2052480.1:2000SEP08	1501410H1	1	208
195	LI:2052480.1:2000SEP08	70846047V1	10	560
195	LI:2052480.1:2000SEP08	71223828V1	904	1354
196	LG:450105.1:2000SEP08	5912415F8	1	376
196	LG:450105.1:2000SEP08	5912415H1	1	299
196	LG:450105.1:2000SEP08	5912415F6	12	565
196	LG:450105.1:2000SEP08	5912415T9	66	535
197	LG:450581.1:2000SEP08	5906909T6	1	410
197	LG:450581.1:2000SEP08	5906909F6	1	501
197	LG:450581.1:2000SEP08	5906909T9	1	432
197	LG:450581.1:2000SEP08	5906909T8	1	359
197	LG:450581.1:2000SEP08	5906909F8	10	484
197	LG:450581.1:2000SEP08	5906909H1	10	302
198	LG:450887.1:2000SEP08	5911592T6	1	523
198	LG:450887.1:2000SEP08	5911592H1	1	290
198	LG:450887.1:2000SEP08	5911592T8	1	473
198	LG:450887.1:2000SEP08	5911592F8	1	569
198	LG:450887.1:2000SEP08	5911592T9	1	473
198	LG:450887.1:2000SEP08	5911592F6	1	565
199	LG:460809.1:2000SEP08	4119207F6	1	336
199	LG:460809.1:2000SEP08	4119207T6	1	336
199	LG:460809.1:2000SEP08	4119207H1	1	175

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
200	LG:452089.1:2000SEP08	5905252H1	35	313
200	LG:452089.1:2000SEP08	5905252T9	236	685
200	LG:452089.1:2000SEP08	5905252T6	376	823
200	LG:452089.1:2000SEP08	5905252F8	1	497
200	LG:452089.1:2000SEP08	5905252F6	35	576
201	LG:1099416.1:2000SEP08	6729842H1	1	412
201	LG:1099416.1:2000SEP08	5401350H1	1	105
201	LG:1099416.1:2000SEP08	6057617H1	56	643
201	LG:1099416.1:2000SEP08	5401350T9	82	666
201	LG:1099416.1:2000SEP08	g3214092	406	782
201	LG:1099416.1:2000SEP08	3524102H1	479	779
202	LG:255713.1:2000SEP08	3439989F6	1	433
202	LG:255713.1:2000SEP08	3439989H1	1	242
202	LG:255713.1:2000SEP08	g1689833	19	402
202	LG:255713.1:2000SEP08	2272896H1	21	157
202	LG:255713.1:2000SEP08	4884069T6	39	468
202	LG:255713.1:2000SEP08	4541467H1	50	115
202	LG:255713.1:2000SEP08	3439989T6	204	480
202	LG:255713.1:2000SEP08	4644925H1	239	467
203	LG:998903.1:2000SEP08	6271004T8	1	378
203	LG:998903.1:2000SEP08	6271004H2	1	277
203	LG:998903.1:2000SEP08	6271004F8	1	604
204	LG:1119656.1:2000SEP08	7741826H1	1	617
204	LG:1119656.1:2000SEP08	4969073H1	26	111
204	LG:1119656.1:2000SEP08	g6199339	69	339
205	LG:1096907.1:2000SEP08	6793205H1	1	353
205	LG:1096907.1:2000SEP08	6938916H1	2	455
205	LG:1096907.1:2000SEP08	6793205F8	17	330
205	LG:1096907.1:2000SEP08	6791360F8	18	350
205	LG:1096907.1:2000SEP08	6798621F8	18	350
205	LG:1096907.1:2000SEP08	6791360T8	18	245
205	LG:1096907.1:2000SEP08	6798621H1	18	350
205	LG:1096907.1:2000SEP08	6791360H1	19	350
205	LG:1096907.1:2000SEP08	6793393F8	21	346
205	LG:1096907.1:2000SEP08	g1667827	30	408
205	LG:1096907.1:2000SEP08	g1667806	398	453
205	LG:1096907.1:2000SEP08	g1798962	25	283
206	LG:1323741.1:2000SEP08	6795323T8	101	459
206	LG:1323741.1:2000SEP08	6790451F8	98	528
206	LG:1323741.1:2000SEP08	6790451T8	98	420
206	LG:1323741.1:2000SEP08	6794426H1	101	532
206	LG:1323741.1:2000SEP08	6796814T8	101	427
206	LG:1323741.1:2000SEP08	6795323F8	101	533
206	LG:1323741.1:2000SEP08	6796814F8	101	577
206	LG:1323741.1:2000SEP08	6795425T8	101	431
206	LG:1323741.1:2000SEP08	6795425F8	101	534
206	LG:1323741.1:2000SEP08	6795323H1	102	505
206	LG:1323741.1:2000SEP08	6795425H1	104	432
206	LG:1323741.1:2000SEP08	6796814H1	104	505

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
206	LG:1323741.1:2000SEP08	6790451H1	108	528
206	LG:1323741.1:2000SEP08	4194768H1	110	429
206	LG:1323741.1:2000SEP08	5220546H1	110	285
206	LG:1323741.1:2000SEP08	g1614987	116	513
206	LG:1323741.1:2000SEP08	5163260H1	122	370
206	LG:1323741.1:2000SEP08	2906282H1	148	391
206	LG:1323741.1:2000SEP08	732838H1	162	243
206	LG:1323741.1:2000SEP08	g1614885	204	545
206	LG:1323741.1:2000SEP08	5794252H1	1	268
206	LG:1323741.1:2000SEP08	5794252F6	1	268
206	LG:1323741.1:2000SEP08	5792265H1	20	268
206	LG:1323741.1:2000SEP08	5790210H1	1	268
206	LG:1323741.1:2000SEP08	5794810H1	1	268
207	LG:1098372.1:2000SEP08	3033193F6	1	272
207	LG:1098372.1:2000SEP08	3033193H1	1	216
207	LG:1098372.1:2000SEP08	5664835T9	31	414
207	LG:1098372.1:2000SEP08	3033193T6	129	443
208	LG:1006783.1:2000SEP08	6792143T8	1	426
208	LG:1006783.1:2000SEP08	6792143F8	1	552
208	LG:1006783.1:2000SEP08	6792143H1	1	532
209	LG:1097562.1:2000SEP08	6796616H1	1	454
209	LG:1097562.1:2000SEP08	6796616T8	1	408
209	LG:1097562.1:2000SEP08	6796616F8	1	510
209	LG:1097562.1:2000SEP08	6798275T8	315	406
209	LG:1097562.1:2000SEP08	6798275H1	1	492
210	LG:998868.1:2000SEP08	6269958F8	1	345
210	LG:998868.1:2000SEP08	6269958H1	1	469
210	LG:998868.1:2000SEP08	6269958T8	293	868
211	LG:1063383.1:2000SEP08	198987H1	1	162
211	LG:1063383.1:2000SEP08	198987R6	1	503
211	LG:1063383.1:2000SEP08	g2885034	702	894
211	LG:1063383.1:2000SEP08	176170H1	12	334
211	LG:1063383.1:2000SEP08	7756883J1	73	671
211	LG:1063383.1:2000SEP08	2531743H1	109	335
211	LG:1063383.1:2000SEP08	146685R1	113	609
211	LG:1063383.1:2000SEP08	146685H1	121	291
211	LG:1063383.1:2000SEP08	4935334H1	213	487
211	LG:1063383.1:2000SEP08	6327776H1	328	623
211	LG:1063383.1:2000SEP08	146685F1	360	1010
211	LG:1063383.1:2000SEP08	198987T6	371	971
211	LG:1063383.1:2000SEP08	1544811H1	458	635
211	LG:1063383.1:2000SEP08	3567613H1	537	852
211	LG:1063383.1:2000SEP08	g3254441	595	1010
211	LG:1063383.1:2000SEP08	2232291H1	617	857
211	LG:1063383.1:2000SEP08	5508171H1	814	1018
212	LG:1400567.1:2000SEP08	7124338H1	138	589
212	LG:1400567.1:2000SEP08	6096304H1	249	408
212	LG:1400567.1:2000SEP08	g5672428	1	366
212	LG:1400567.1:2000SEP08	g6473179	38	363

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
213	U:449404.1:2000SEP08	5908301H1	1	311
213	U:449404.1:2000SEP08	5908301F8	1	519
213	U:449404.1:2000SEP08	6271267F8	24	643
213	U:449404.1:2000SEP08	6271267H2	24	492
213	U:449404.1:2000SEP08	5908301T9	248	582
214	U:449941.2:2000SEP08	5911361F8	1	206
214	U:449941.2:2000SEP08	5911361H1	1	301
214	U:449941.2:2000SEP08	5911361T8	238	780
214	U:449941.2:2000SEP08	5911361T9	249	710
215	U:450229.1:2000SEP08	6268506H1	1	474
215	U:450229.1:2000SEP08	5913065F7	1	509
215	U:450229.1:2000SEP08	6268506F8	1	589
215	U:450229.1:2000SEP08	5913065F8	1	567
215	U:450229.1:2000SEP08	6268506T8	1	600
215	U:450229.1:2000SEP08	5913065H1	7	303
215	U:450229.1:2000SEP08	5913065T7	118	715
216	U:450399.3:2000SEP08	71464335V1	1	627
216	U:450399.3:2000SEP08	71452422V1	1	248
216	U:450399.3:2000SEP08	71457485V1	1	615
216	U:450399.3:2000SEP08	71465203V1	1	612
216	U:450399.3:2000SEP08	71465621V1	1	603
216	U:450399.3:2000SEP08	71467288V1	1	526
216	U:450399.3:2000SEP08	5910662F8	1	510
216	U:450399.3:2000SEP08	71449251V1	1	232
216	U:450399.3:2000SEP08	71461359V1	1	488
216	U:450399.3:2000SEP08	71457006V1	1	318
216	U:450399.3:2000SEP08	71465838V1	1	511
216	U:450399.3:2000SEP08	71440536V1	1	405
216	U:450399.3:2000SEP08	71455851V1	1	248
216	U:450399.3:2000SEP08	5910662F6	1	552
216	U:450399.3:2000SEP08	71455734V1	1	143
216	U:450399.3:2000SEP08	5910662H1	1	295
216	U:450399.3:2000SEP08	5910662T6	9	601
216	U:450399.3:2000SEP08	71464857V1	17	469
216	U:450399.3:2000SEP08	71458424V1	53	155
216	U:450399.3:2000SEP08	5910662T9	67	572
216	U:450399.3:2000SEP08	5910662T8	146	589
216	U:450399.3:2000SEP08	71439746V1	173	327
216	U:450399.3:2000SEP08	71437264V1	205	476
216	U:450399.3:2000SEP08	71439082V1	353	504
216	U:450399.3:2000SEP08	71463937V1	367	571
216	U:450399.3:2000SEP08	71459908V1	430	693
217	U:455771.1:2000SEP08	5911540F8	1	460
217	U:455771.1:2000SEP08	5911540H1	1	250
217	U:455771.1:2000SEP08	5911540T9	27	568
217	U:455771.1:2000SEP08	5911540T8	78	569
218	U:720459.1:2000SEP08	8081371H2	1	347
218	U:720459.1:2000SEP08	6571208H1	1	472
218	U:720459.1:2000SEP08	6571208F8	1	489

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
218	LI:720459.1:2000SEP08	6571208T8	11	564
219	LI:723156.1:2000SEP08	6270651T8	1	429
219	LI:723156.1:2000SEP08	6270651H2	26	533
219	LI:723156.1:2000SEP08	6270651F8	26	508
220	LI:728055.1:2000SEP08	3565507T6	98	504
220	LI:728055.1:2000SEP08	6793060H1	1	476
221	LI:1020789.1:2000SEP08	6794842H1	1	442
221	LI:1020789.1:2000SEP08	6794842T8	1	405
221	LI:1020789.1:2000SEP08	6794842F8	1	316
222	LI:1071728.1:2000SEP08	6790925T8	1	375
222	LI:1071728.1:2000SEP08	6792395T8	1	380
222	LI:1071728.1:2000SEP08	6792395F8	1	498
222	LI:1071728.1:2000SEP08	6792395H1	1	487
222	LI:1071728.1:2000SEP08	6790925F8	1	482
222	LI:1071728.1:2000SEP08	6791755T8	1	256
222	LI:1071728.1:2000SEP08	6791755H1	1	485
222	LI:1071728.1:2000SEP08	6790925H1	1	163
222	LI:1071728.1:2000SEP08	6791755F8	1	482
223	LI:1084329.1:2000SEP08	6791107H1	1	256
223	LI:1084329.1:2000SEP08	6797814F8	1	498
223	LI:1084329.1:2000SEP08	6797814H1	1	498
223	LI:1084329.1:2000SEP08	6791987F8	1	496
223	LI:1084329.1:2000SEP08	6791987T8	1	394
223	LI:1084329.1:2000SEP08	6791987H1	1	496
223	LI:1084329.1:2000SEP08	6791107F8	12	522
223	LI:1084329.1:2000SEP08	6798578F8	82	500
224	LI:246422.1:2000SEP08	71544580V1	1	429
224	LI:246422.1:2000SEP08	71546409V1	1	413
224	LI:246422.1:2000SEP08	3033193F6	1	272
224	LI:246422.1:2000SEP08	3033193H1	1	216
224	LI:246422.1:2000SEP08	5664835T9	31	413
224	LI:246422.1:2000SEP08	3033193T6	129	442
224	LI:246422.1:2000SEP08	71544508V1	133	413
224	LI:246422.1:2000SEP08	71543126V1	152	302
224	LI:246422.1:2000SEP08	71545653V1	175	413
224	LI:246422.1:2000SEP08	71543870V1	318	439
224	LI:246422.1:2000SEP08	71545652V1	357	413
225	LI:1086066.1:2000SEP08	6793877T8	194	601
225	LI:1086066.1:2000SEP08	6793877H1	1	604
226	LI:223142.1:2000SEP08	797887R6	222	612
226	LI:223142.1:2000SEP08	5377283T9	59	591
226	LI:223142.1:2000SEP08	g4078031	284	558
226	LI:223142.1:2000SEP08	797887H1	222	430
226	LI:223142.1:2000SEP08	5377283H1	1	231
226	LI:223142.1:2000SEP08	5443396T9	177	700
226	LI:223142.1:2000SEP08	g5675714	214	679
226	LI:223142.1:2000SEP08	g1313047	216	648
226	LI:223142.1:2000SEP08	5443396H1	1068	1207
226	LI:223142.1:2000SEP08	5443396F9	694	1207



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
226	LI:223142.1:2000SEP08	5443396F8	726	1205
226	LI:223142.1:2000SEP08	g3424259	705	968
226	LI:223142.1:2000SEP08	g2616031	723	931
226	LI:223142.1:2000SEP08	g4089713	477	929
226	LI:223142.1:2000SEP08	g5849958	673	922
226	LI:223142.1:2000SEP08	4642964T6	347	886
226	LI:223142.1:2000SEP08	797887T6	294	883
226	LI:223142.1:2000SEP08	7431990H1	256	855
226	LI:223142.1:2000SEP08	5377283F8	20	125
227	LI:885368.1:2000SEP08	5909782T7	1	492
227	LI:885368.1:2000SEP08	6269652H1	1	413
227	LI:885368.1:2000SEP08	6269652F8	1	680
227	LI:885368.1:2000SEP08	5909782F7	1	576
227	LI:885368.1:2000SEP08	5909895H1	1	133
227	LI:885368.1:2000SEP08	5909795H1	1	303
227	LI:885368.1:2000SEP08	5909782H1	1	304
227	LI:885368.1:2000SEP08	6269652T8	99	616
227	LI:885368.1:2000SEP08	4982679T6	375	548
228	LI:481782.1:2000SEP08	5905477F6	1	566
228	LI:481782.1:2000SEP08	5905477H1	1	272
228	LI:481782.1:2000SEP08	5912506F6	16	232
228	LI:481782.1:2000SEP08	5912506F8	35	371
228	LI:481782.1:2000SEP08	5912506H1	35	297
228	LI:481782.1:2000SEP08	4008278H1	217	352
228	LI:481782.1:2000SEP08	5905477T9	413	934
229	LI:1093813.1:2000SEP08	6790451F8	1	431
229	LI:1093813.1:2000SEP08	6790451T8	1	323
229	LI:1093813.1:2000SEP08	6794426H1	4	435
229	LI:1093813.1:2000SEP08	6795425F8	4	437
229	LI:1093813.1:2000SEP08	6795323T8	4	362
229	LI:1093813.1:2000SEP08	6795323F8	4	436
229	LI:1093813.1:2000SEP08	6795425T8	4	334
229	LI:1093813.1:2000SEP08	6796814F8	4	480
229	LI:1093813.1:2000SEP08	6796814T8	4	330
229	LI:1093813.1:2000SEP08	6795323H1	5	408
229	LI:1093813.1:2000SEP08	6796814H1	7	408
229	LI:1093813.1:2000SEP08	6795425H1	7	335
229	LI:1093813.1:2000SEP08	6790451H1	4	431
230	LI:449413.2:2000SEP08	6271008T8	1	528
230	LI:449413.2:2000SEP08	6271008H2	1	470
230	LI:449413.2:2000SEP08	6271008F8	1	543
231	LI:450105.1:2000SEP08	5912415H1	1	301
231	LI:450105.1:2000SEP08	5912415F6	12	567
231	LI:450105.1:2000SEP08	5912415T9	66	537
231	LI:450105.1:2000SEP08	5912415F8	1	378
232	LI:814285.1:2000SEP08	71499034V1	1	213
232	LI:814285.1:2000SEP08	71499323V1	1	213
232	LI:814285.1:2000SEP08	71681849V1	1	171
232	LI:814285.1:2000SEP08	71684640V1	1	171

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
232	LI:814285.1:2000SEP08	71530104V1	1	135
232	LI:814285.1:2000SEP08	71529064V1	1	122
232	LI:814285.1:2000SEP08	3966795H1	1	290
232	LI:814285.1:2000SEP08	3274864H1	14	288
232	LI:814285.1:2000SEP08	71683688V1	1	629
232	LI:814285.1:2000SEP08	71681753V1	1	564
232	LI:814285.1:2000SEP08	71685121V1	1	521
232	LI:814285.1:2000SEP08	71687268V1	336	764
232	LI:814285.1:2000SEP08	71687355V1	1	752
232	LI:814285.1:2000SEP08	3966795T6	24	703
232	LI:814285.1:2000SEP08	71682443V1	157	681
232	LI:814285.1:2000SEP08	71683031V1	1	431
232	LI:814285.1:2000SEP08	3632915H1	259	517
232	LI:814285.1:2000SEP08	71683429V1	1	468
232	LI:814285.1:2000SEP08	71685862V1	1	459
232	LI:814285.1:2000SEP08	71683642V1	1	439
232	LI:814285.1:2000SEP08	71471770V1	1	374
232	LI:814285.1:2000SEP08	71498058V1	1	290
232	LI:814285.1:2000SEP08	71685989V1	1	396
232	LI:814285.1:2000SEP08	3966795F6	1	397
233	LI:1142855.1:2000SEP08	7444943T2	1	522
233	LI:1142855.1:2000SEP08	4630066H1	198	459
233	LI:1142855.1:2000SEP08	5941205H1	1	162
234	LI:817330.1:2000SEP08	6271004F8	10	624
234	LI:817330.1:2000SEP08	6271004H2	9	294
234	LI:817330.1:2000SEP08	6271004T8	1	398
235	LI:817845.1:2000SEP08	6268754H1	1	568
235	LI:817845.1:2000SEP08	5906668H1	1	253
235	LI:817845.1:2000SEP08	6268754T8	1	599
235	LI:817845.1:2000SEP08	6268754F8	21	687
236	LI:460809.1:2000SEP08	5848081F8	1	196
236	LI:460809.1:2000SEP08	5848081H1	1	149
236	LI:460809.1:2000SEP08	4119207F6	71	406
236	LI:460809.1:2000SEP08	4119207T6	71	406
236	LI:460809.1:2000SEP08	4119207H1	71	245
237	LI:815874.1:2000SEP08	55000166T1	1	463
237	LI:815874.1:2000SEP08	55000166T2	1	463
237	LI:815874.1:2000SEP08	55002307J1	47	566
237	LI:815874.1:2000SEP08	4577588H1	48	292
237	LI:815874.1:2000SEP08	5473293H1	48	198
237	LI:815874.1:2000SEP08	7626863J1	50	721
237	LI:815874.1:2000SEP08	6215083H1	132	628
237	LI:815874.1:2000SEP08	6215083T8	132	523
237	LI:815874.1:2000SEP08	6215083F8	144	626
238	LI:255713.1:2000SEP08	71631750V1	1	537
238	LI:255713.1:2000SEP08	3439989F6	1	436
238	LI:255713.1:2000SEP08	3439989H1	1	243
238	LI:255713.1:2000SEP08	71627522V1	1	98
238	LI:255713.1:2000SEP08	71630884V1	1	489

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
238	LI:255713.1:2000SEP08	71631310V1	1	527
238	LI:255713.1:2000SEP08	g1689833	19	405
238	LI:255713.1:2000SEP08	2272896H1	21	157
238	LI:255713.1:2000SEP08	4884069T6	39	471
238	LI:255713.1:2000SEP08	4541467H1	50	115
238	LI:255713.1:2000SEP08	3733119T6	113	603
238	LI:255713.1:2000SEP08	3439989T6	205	490
238	LI:255713.1:2000SEP08	4644925H1	241	478
238	LI:255713.1:2000SEP08	g3118716	492	817
238	LI:255713.1:2000SEP08	3733119F6	578	918
238	LI:255713.1:2000SEP08	3733119H1	650	918
239	LI:035973.1:2000SEP08	5401350T9	83	680
239	LI:035973.1:2000SEP08	6057617H1	56	657
239	LI:035973.1:2000SEP08	7671773H1	811	946
239	LI:035973.1:2000SEP08	7671748H1	381	938
239	LI:035973.1:2000SEP08	g3214092	407	796
239	LI:035973.1:2000SEP08	3524102H1	488	793
239	LI:035973.1:2000SEP08	6729842H1	1	413
239	LI:035973.1:2000SEP08	5401350H1	1	106
240	LI:1138110.1:2000SEP08	7741826H1	1	618
240	LI:1138110.1:2000SEP08	4969073H1	26	112
240	LI:1138110.1:2000SEP08	g6199339	69	340
241	LI:2049074.1:2000SEP08	6792402T8	9	607
241	LI:2049074.1:2000SEP08	6792402F8	9	720
241	LI:2049074.1:2000SEP08	6792402H1	9	515
241	LI:2049074.1:2000SEP08	6798273T8	368	558
241	LI:2049074.1:2000SEP08	6798273F8	1	589
241	LI:2049074.1:2000SEP08	6798273H1	1	512
241	LI:2049074.1:2000SEP08	4431872H1	3	85
242	LI:1092460.1:2000SEP08	6791420H1	1	207
242	LI:1092460.1:2000SEP08	6793081T8	1	560
242	LI:1092460.1:2000SEP08	6791420F8	1	284
242	LI:1092460.1:2000SEP08	6791420T8	1	556
242	LI:1092460.1:2000SEP08	6793081F8	1	670
242	LI:1092460.1:2000SEP08	6793081H1	1	437
243	LI:399421.1:2000SEP08	71674115V1	768	1316
243	LI:399421.1:2000SEP08	71673332V1	386	890
243	LI:399421.1:2000SEP08	71671508V1	462	1005
243	LI:399421.1:2000SEP08	71671885V1	462	1005
243	LI:399421.1:2000SEP08	71673708V1	575	1229
243	LI:399421.1:2000SEP08	3532689H1	2	272
243	LI:399421.1:2000SEP08	g5671055	14	474
243	LI:399421.1:2000SEP08	71545810V1	827	1065
243	LI:399421.1:2000SEP08	71673382V1	1560	2213
243	LI:399421.1:2000SEP08	71674031V1	1563	2178
243	LI:399421.1:2000SEP08	71674569V1	1563	1994
243	LI:399421.1:2000SEP08	71671427V1	1564	2145
243	LI:399421.1:2000SEP08	71671434V1	1579	1674
243	LI:399421.1:2000SEP08	71671021V1	1582	2109

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
243	LI:399421.1:2000SEP08	71672327V1	1651	2327
243	LI:399421.1:2000SEP08	71671131V1	1685	2101
243	LI:399421.1:2000SEP08	71670034V1	1810	2552
243	LI:399421.1:2000SEP08	71669677V1	1860	2565
243	LI:399421.1:2000SEP08	71669977V1	1868	2471
243	LI:399421.1:2000SEP08	71669922V1	1868	2262
243	LI:399421.1:2000SEP08	71670815V1	1868	2174
243	LI:399421.1:2000SEP08	71674026V1	2090	2527
243	LI:399421.1:2000SEP08	71672887V1	2102	2577
243	LI:399421.1:2000SEP08	71669725V1	1437	1936
243	LI:399421.1:2000SEP08	71671316V1	1451	2129
243	LI:399421.1:2000SEP08	71673713V1	1469	2054
243	LI:399421.1:2000SEP08	71672796V1	1	665
243	LI:399421.1:2000SEP08	71672127V1	52	714
243	LI:399421.1:2000SEP08	g7455758	83	458
243	LI:399421.1:2000SEP08	g2457980	96	469
243	LI:399421.1:2000SEP08	71672645V1	167	936
243	LI:399421.1:2000SEP08	71671802V1	194	796
243	LI:399421.1:2000SEP08	71674174V1	245	1000
243	LI:399421.1:2000SEP08	71671116V1	244	1036
243	LI:399421.1:2000SEP08	71672884V1	286	931
243	LI:399421.1:2000SEP08	71672574V1	323	1079
243	LI:399421.1:2000SEP08	71671331V1	332	984
243	LI:399421.1:2000SEP08	71672409V1	338	774
243	LI:399421.1:2000SEP08	g2063420	349	598
243	LI:399421.1:2000SEP08	71673338V1	356	711
243	LI:399421.1:2000SEP08	71542839V1	428	896
243	LI:399421.1:2000SEP08	71674325V1	1175	1690
243	LI:399421.1:2000SEP08	71670040V1	1194	1691
243	LI:399421.1:2000SEP08	71670586V1	1205	1690
243	LI:399421.1:2000SEP08	71672111V1	1244	1929
243	LI:399421.1:2000SEP08	71670467V1	1282	1691
243	LI:399421.1:2000SEP08	71670695V1	1284	1690
243	LI:399421.1:2000SEP08	71670674V1	1283	1648
243	LI:399421.1:2000SEP08	71675221V1	1314	1959
243	LI:399421.1:2000SEP08	g3734807	1316	1763
243	LI:399421.1:2000SEP08	71670506V1	1327	2017
243	LI:399421.1:2000SEP08	71673468V1	1331	1674
243	LI:399421.1:2000SEP08	71674268V1	1331	1674
243	LI:399421.1:2000SEP08	71670681V1	1343	1690
243	LI:399421.1:2000SEP08	71674973V1	1363	1690
243	LI:399421.1:2000SEP08	71671128V1	1369	2118
243	LI:399421.1:2000SEP08	71554925V1	1375	1661
243	LI:399421.1:2000SEP08	71674341V1	1420	2112
243	LI:399421.1:2000SEP08	71673770V1	1433	2043
243	LI:399421.1:2000SEP08	71675202V1	819	1516
243	LI:399421.1:2000SEP08	71540274V1	1169	1661
243	LI:399421.1:2000SEP08	71670585V1	1159	1805
243	LI:399421.1:2000SEP08	71674048V1	1174	1679

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
243	LI:399421.1:2000SEP08	71673538V1	841	1397
243	LI:399421.1:2000SEP08	71673548V1	886	1614
243	LI:399421.1:2000SEP08	71675105V1	900	1584
243	LI:399421.1:2000SEP08	71670075V1	906	1546
243	LI:399421.1:2000SEP08	71671772V1	919	1677
243	LI:399421.1:2000SEP08	71671422V1	936	1669
243	LI:399421.1:2000SEP08	71674671V1	762	1357
243	LI:399421.1:2000SEP08	71672673V1	982	1317
243	LI:399421.1:2000SEP08	71672580V1	1086	1816
243	LI:399421.1:2000SEP08	71671338V1	1091	1690
243	LI:399421.1:2000SEP08	71670081V1	580	1262
243	LI:399421.1:2000SEP08	3532689F6	1	381
243	LI:399421.1:2000SEP08	71671278V1	1	650
243	LI:399421.1:2000SEP08	71673438V1	356	890
243	LI:399421.1:2000SEP08	71675265V1	782	1353
243	LI:399421.1:2000SEP08	71674149V1	1141	1657
243	LI:399421.1:2000SEP08	71541507V1	1169	1669
243	LI:399421.1:2000SEP08	71670447V1	1179	1936
243	LI:399421.1:2000SEP08	71671453V1	982	1317
243	LI:399421.1:2000SEP08	71671269V1	1082	1690
243	LI:399421.1:2000SEP08	71669745V1	381	997
243	LI:399421.1:2000SEP08	71672102V1	1	613
243	LI:399421.1:2000SEP08	71673585V1	362	1095
243	LI:399421.1:2000SEP08	71673905V1	789	1516
243	LI:399421.1:2000SEP08	71670267V1	445	1133
243	LI:399421.1:2000SEP08	71671650V1	1028	1753
243	LI:399421.1:2000SEP08	71670603V1	1027	1496
243	LI:399421.1:2000SEP08	71674515V1	1147	1560
243	LI:399421.1:2000SEP08	71541489V1	1150	1691
243	LI:399421.1:2000SEP08	71674543V1	1051	1694
243	LI:399421.1:2000SEP08	71674611V1	1077	1456
243	LI:399421.1:2000SEP08	71673288V1	399	1008
243	LI:399421.1:2000SEP08	71672026V1	573	1294
243	LI:399421.1:2000SEP08	71538613V1	583	757
243	LI:399421.1:2000SEP08	71673555V1	610	1096
243	LI:399421.1:2000SEP08	71671221V1	649	1367
243	LI:399421.1:2000SEP08	71671332V1	675	1202
243	LI:399421.1:2000SEP08	71548524V1	703	1054
243	LI:399421.1:2000SEP08	71670624V1	721	1063
243	LI:399421.1:2000SEP08	71673372V1	724	1359
243	LI:399421.1:2000SEP08	71673387V1	723	1127
243	LI:399421.1:2000SEP08	71671687V1	751	1342
243	LI:399421.1:2000SEP08	71669849V1	1109	1690
243	LI:399421.1:2000SEP08	g2458468	25	469
243	LI:399421.1:2000SEP08	g6029855	14	483
243	LI:399421.1:2000SEP08	g3785048	14	329
243	LI:399421.1:2000SEP08	71547880V1	23	310
244	LI:816655.2:2000SEP08	2307022H1	999	1182
244	LI:816655.2:2000SEP08	g6074114	1000	1285

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
244	LI:816655.2:2000SEP08	g4069167	775	1286
244	LI:816655.2:2000SEP08	995379H1	776	1142
244	LI:816655.2:2000SEP08	g2270223	777	1285
244	LI:816655.2:2000SEP08	g3884489	833	1287
244	LI:816655.2:2000SEP08	2274905H1	754	1053
244	LI:816655.2:2000SEP08	603456H1	753	1032
244	LI:816655.2:2000SEP08	g5553619	867	1285
244	LI:816655.2:2000SEP08	g1376743	869	1193
244	LI:816655.2:2000SEP08	g4137994	881	1285
244	LI:816655.2:2000SEP08	5594093H1	879	1164
244	LI:816655.2:2000SEP08	g766356	885	1034
244	LI:816655.2:2000SEP08	2197723H1	884	1172
244	LI:816655.2:2000SEP08	1581140H1	885	1118
244	LI:816655.2:2000SEP08	2999223H1	891	1216
244	LI:816655.2:2000SEP08	1460549H1	892	1156
244	LI:816655.2:2000SEP08	645746H1	904	1193
244	LI:816655.2:2000SEP08	791389H1	904	1049
244	LI:816655.2:2000SEP08	g6034220	906	1286
244	LI:816655.2:2000SEP08	g6029653	907	1279
244	LI:816655.2:2000SEP08	6177607H1	908	1237
244	LI:816655.2:2000SEP08	1927953H1	908	1202
244	LI:816655.2:2000SEP08	562727H1	910	1149
244	LI:816655.2:2000SEP08	g6133735	915	1285
244	LI:816655.2:2000SEP08	2923501H1	920	1234
244	LI:816655.2:2000SEP08	4775589H1	920	1226
244	LI:816655.2:2000SEP08	5261172H1	920	1205
244	LI:816655.2:2000SEP08	1541796H1	920	1165
244	LI:816655.2:2000SEP08	3146255H1	926	1244
244	LI:816655.2:2000SEP08	2213394H1	925	1194
244	LI:816655.2:2000SEP08	4322891H1	925	1205
244	LI:816655.2:2000SEP08	1311827H1	929	1209
244	LI:816655.2:2000SEP08	g2348484	758	1285
244	LI:816655.2:2000SEP08	4510035H1	757	1072
244	LI:816655.2:2000SEP08	g2410627	759	1285
244	LI:816655.2:2000SEP08	g1377471	759	1179
244	LI:816655.2:2000SEP08	g3255360	831	1285
244	LI:816655.2:2000SEP08	g3038648	831	1285
244	LI:816655.2:2000SEP08	g2656902	831	1285
244	LI:816655.2:2000SEP08	g6198637	832	1292
244	LI:816655.2:2000SEP08	g1382354	834	1286
244	LI:816655.2:2000SEP08	g4890713	832	1283
244	LI:816655.2:2000SEP08	g2336858	834	1285
244	LI:816655.2:2000SEP08	387010H1	844	1148
244	LI:816655.2:2000SEP08	70507819D1	835	1285
244	LI:816655.2:2000SEP08	3662763H1	835	1156
244	LI:816655.2:2000SEP08	4593659H1	841	1158
244	LI:816655.2:2000SEP08	3695060H1	845	1158
244	LI:816655.2:2000SEP08	6715881H1	859	1285
244	LI:816655.2:2000SEP08	4880202H1	859	1159

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
244	LI:816655.2:2000SEP08	815968H1	859	1134
244	LI:816655.2:2000SEP08	1346582H1	859	1081
244	LI:816655.2:2000SEP08	g6400352	864	1285
244	LI:816655.2:2000SEP08	1914372H1	864	1146
244	LI:816655.2:2000SEP08	g4619759	846	1285
244	LI:816655.2:2000SEP08	g6439582	848	1285
244	LI:816655.2:2000SEP08	5326753H1	848	1142
244	LI:816655.2:2000SEP08	g6438934	851	1285
244	LI:816655.2:2000SEP08	g5658518	852	1285
244	LI:816655.2:2000SEP08	1422529H1	853	1104
244	LI:816655.2:2000SEP08	g6196473	855	1285
244	LI:816655.2:2000SEP08	5435388H1	857	1099
244	LI:816655.2:2000SEP08	2647919H1	858	1140
244	LI:816655.2:2000SEP08	1384031H1	867	1138
244	LI:816655.2:2000SEP08	4371871H1	858	1127
244	LI:816655.2:2000SEP08	5566646H1	858	1133
244	LI:816655.2:2000SEP08	2925562H1	724	1032
244	LI:816655.2:2000SEP08	3121148H1	725	1088
244	LI:816655.2:2000SEP08	970837H1	725	1070
244	LI:816655.2:2000SEP08	3121685H1	725	1052
244	LI:816655.2:2000SEP08	2653956H1	777	1119
244	LI:816655.2:2000SEP08	1858090H1	822	1135
244	LI:816655.2:2000SEP08	g1547620	822	1131
244	LI:816655.2:2000SEP08	670081H1	822	1122
244	LI:816655.2:2000SEP08	g5425145	825	1285
244	LI:816655.2:2000SEP08	1981888T6	823	1237
244	LI:816655.2:2000SEP08	g3422463	824	1286
244	LI:816655.2:2000SEP08	g1422803	826	1287
244	LI:816655.2:2000SEP08	g3694133	825	1279
244	LI:816655.2:2000SEP08	g2110622	826	1286
244	LI:816655.2:2000SEP08	g3897973	826	1285
244	LI:816655.2:2000SEP08	g1224694	790	1285
244	LI:816655.2:2000SEP08	70247426V1	1	477
244	LI:816655.2:2000SEP08	70250858V1	1	443
244	LI:816655.2:2000SEP08	7690215H1	1	354
244	LI:816655.2:2000SEP08	839665H1	1	147
244	LI:816655.2:2000SEP08	133847H1	1	173
244	LI:816655.2:2000SEP08	6408445H1	726	1285
244	LI:816655.2:2000SEP08	3111675H1	726	1048
244	LI:816655.2:2000SEP08	g4437633	772	1285
244	LI:816655.2:2000SEP08	g3664889	792	1279
244	LI:816655.2:2000SEP08	g5848288	792	1284
244	LI:816655.2:2000SEP08	g5431438	795	1285
244	LI:816655.2:2000SEP08	2937993H1	796	1110
244	LI:816655.2:2000SEP08	70444144D1	798	1237
244	LI:816655.2:2000SEP08	607326H1	808	1082
244	LI:816655.2:2000SEP08	g6047082	800	1285
244	LI:816655.2:2000SEP08	1860615F6	800	1279
244	LI:816655.2:2000SEP08	1860606H1	800	1159

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
244	LI:816655.2:2000SEP08	4349840H1	808	1086
244	LI:816655.2:2000SEP08	4323392H1	800	1103
244	LI:816655.2:2000SEP08	5592702H1	800	1107
244	LI:816655.2:2000SEP08	563915H1	800	1059
244	LI:816655.2:2000SEP08	3256662H1	800	1085
244	LI:816655.2:2000SEP08	g3255332	801	1277
244	LI:816655.2:2000SEP08	g1578030	801	1126
244	LI:816655.2:2000SEP08	60110954B2	802	1271
244	LI:816655.2:2000SEP08	4881781H1	806	1022
244	LI:816655.2:2000SEP08	3444053H1	807	1135
244	LI:816655.2:2000SEP08	3174634H1	807	1072
244	LI:816655.2:2000SEP08	6299502H1	808	1149
244	LI:816655.2:2000SEP08	7360181H1	608	1204
244	LI:816655.2:2000SEP08	6620724H1	623	1222
244	LI:816655.2:2000SEP08	70507337D1	788	1285
244	LI:816655.2:2000SEP08	g4988595	790	1285
244	LI:816655.2:2000SEP08	3942547T6	792	1258
244	LI:816655.2:2000SEP08	6517011H1	726	1286
244	LI:816655.2:2000SEP08	6549137H1	724	1204
244	LI:816655.2:2000SEP08	70995774V1	713	1284
244	LI:816655.2:2000SEP08	70444578D1	714	1127
244	LI:816655.2:2000SEP08	1602631T6	717	1202
244	LI:816655.2:2000SEP08	70442740D1	721	1265
244	LI:816655.2:2000SEP08	g6228483	778	1285
244	LI:816655.2:2000SEP08	g3330557	780	1285
244	LI:816655.2:2000SEP08	70248539V1	446	966
244	LI:816655.2:2000SEP08	g1337029	665	1200
244	LI:816655.2:2000SEP08	6512274H1	606	1203
244	LI:816655.2:2000SEP08	2838160H2	606	894
244	LI:816655.2:2000SEP08	5267260H1	978	1209
244	LI:816655.2:2000SEP08	4651749H1	969	1209
244	LI:816655.2:2000SEP08	3326922H1	16	299
244	LI:816655.2:2000SEP08	g4137208	830	1279
244	LI:816655.2:2000SEP08	g3037949	829	1286
244	LI:816655.2:2000SEP08	1219349H1	829	1077
244	LI:816655.2:2000SEP08	5511557H1	725	972
244	LI:816655.2:2000SEP08	1652441H1	769	1055
244	LI:816655.2:2000SEP08	71435812V1	853	1281
244	LI:816655.2:2000SEP08	g892394	814	1186
244	LI:816655.2:2000SEP08	g2320997	815	1285
244	LI:816655.2:2000SEP08	g2242380	816	1287
244	LI:816655.2:2000SEP08	g6196328	817	1285
244	LI:816655.2:2000SEP08	g2504947	816	1278
244	LI:816655.2:2000SEP08	g5743160	817	1285
244	LI:816655.2:2000SEP08	g3173656	818	1183
244	LI:816655.2:2000SEP08	4349232H1	818	1100
244	LI:816655.2:2000SEP08	g5398329	819	1285
244	LI:816655.2:2000SEP08	g1843836	819	1205
244	LI:816655.2:2000SEP08	2435702H1	819	1107



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
244	LI:816655.2:2000SEP08	g2779404	821	1286
244	LI:816655.2:2000SEP08	g4095635	821	1286
244	LI:816655.2:2000SEP08	4304173H1	821	1129
244	LI:816655.2:2000SEP08	g4088858	822	1290
244	LI:816655.2:2000SEP08	g5398034	834	1285
244	LI:816655.2:2000SEP08	2939062H1	725	1030
244	LI:816655.2:2000SEP08	1602631H1	1	191
244	LI:816655.2:2000SEP08	g6464173	970	1285
244	LI:816655.2:2000SEP08	4606258H1	983	1251
244	LI:816655.2:2000SEP08	1212531H1	975	1206
244	LI:816655.2:2000SEP08	g1376690	975	1157
244	LI:816655.2:2000SEP08	g6464528	976	1285
244	LI:816655.2:2000SEP08	833443H1	990	1054
244	LI:816655.2:2000SEP08	1292202H1	981	1183
244	LI:816655.2:2000SEP08	g1448481	990	1205
244	LI:816655.2:2000SEP08	7128006H1	992	1299
244	LI:816655.2:2000SEP08	70998240V1	508	1123
244	LI:816655.2:2000SEP08	70995534V1	513	1115
244	LI:816655.2:2000SEP08	70996937V1	509	1181
244	LI:816655.2:2000SEP08	70248846V1	513	1025
244	LI:816655.2:2000SEP08	g3415495	792	1285
244	LI:816655.2:2000SEP08	292307H1	791	1139
244	LI:816655.2:2000SEP08	2994979H1	776	1138
244	LI:816655.2:2000SEP08	2227862H1	777	1076
244	LI:816655.2:2000SEP08	4513074H1	777	1071
244	LI:816655.2:2000SEP08	g2806970	778	1286
244	LI:816655.2:2000SEP08	70996274V1	632	1283
244	LI:816655.2:2000SEP08	g2558287	634	1087
244	LI:816655.2:2000SEP08	3318870H1	749	1054
244	LI:816655.2:2000SEP08	6747339H1	562	1193
244	LI:816655.2:2000SEP08	5771213H1	562	1059
244	LI:816655.2:2000SEP08	1215646H1	562	839
244	LI:816655.2:2000SEP08	4610433H1	562	838
244	LI:816655.2:2000SEP08	7084557H1	569	1171
244	LI:816655.2:2000SEP08	5444485F8	575	1158
244	LI:816655.2:2000SEP08	5444485F9	576	1266
244	LI:816655.2:2000SEP08	6721894H1	590	1135
244	LI:816655.2:2000SEP08	5716571H1	629	1140
244	LI:816655.2:2000SEP08	1528653H1	933	1160
244	LI:816655.2:2000SEP08	3239517H1	937	1226
244	LI:816655.2:2000SEP08	1666325H1	941	1180
244	LI:816655.2:2000SEP08	574249H1	948	1217
244	LI:816655.2:2000SEP08	1357273H1	947	1099
244	LI:816655.2:2000SEP08	1344137H1	954	1234
244	LI:816655.2:2000SEP08	4698462H1	954	1224
244	LI:816655.2:2000SEP08	g4109989	965	1285
244	LI:816655.2:2000SEP08	g4522919	963	1285
244	LI:816655.2:2000SEP08	71298721V1	644	1221
244	LI:816655.2:2000SEP08	032779H1	741	1055

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
244	LI:816655.2:2000SEP08	6271389H2	130	688
244	LI:816655.2:2000SEP08	2800611H1	749	1037
244	LI:816655.2:2000SEP08	3881645H1	749	1143
244	LI:816655.2:2000SEP08	g3769757	751	1257
244	LI:816655.2:2000SEP08	1980337H1	752	1054
244	LI:816655.2:2000SEP08	1328165T6	602	1248
244	LI:816655.2:2000SEP08	70998215V1	596	1079
244	LI:816655.2:2000SEP08	6217535H1	593	1121
244	LI:816655.2:2000SEP08	5020927T1	653	1251
244	LI:816655.2:2000SEP08	g4267519	773	1285
244	LI:816655.2:2000SEP08	g3869742	774	1285
244	LI:816655.2:2000SEP08	1951546H1	770	1035
244	LI:816655.2:2000SEP08	3867575H1	771	1090
244	LI:816655.2:2000SEP08	70444345D1	695	1127
244	LI:816655.2:2000SEP08	5444485T9	711	1165
244	LI:816655.2:2000SEP08	70998184V1	748	1285
244	LI:816655.2:2000SEP08	1720422H1	993	1224
244	LI:816655.2:2000SEP08	g3677334	788	1289
244	LI:816655.2:2000SEP08	g5128138	785	1285
244	LI:816655.2:2000SEP08	g3146055	785	1285
244	LI:816655.2:2000SEP08	g2466728	783	1255
244	LI:816655.2:2000SEP08	g3161904	784	1287
244	LI:816655.2:2000SEP08	70995686V1	520	1216
244	LI:816655.2:2000SEP08	5826417H1	560	1182
244	LI:816655.2:2000SEP08	70995343V1	710	1388
244	LI:816655.2:2000SEP08	60111037B2	686	1253
244	LI:816655.2:2000SEP08	897485H1	728	1060
244	LI:816655.2:2000SEP08	2263535H1	5	173
244	LI:816655.2:2000SEP08	g4888169	1070	1285
244	LI:816655.2:2000SEP08	g5634856	1108	1279
244	LI:816655.2:2000SEP08	g6030854	1114	1286
244	LI:816655.2:2000SEP08	g6073520	1131	1285
244	LI:816655.2:2000SEP08	g4989988	1136	1285
244	LI:816655.2:2000SEP08	g4619083	1151	1285
244	LI:816655.2:2000SEP08	4857252H1	781	1113
244	LI:816655.2:2000SEP08	g5596053	810	1279
244	LI:816655.2:2000SEP08	3788678H1	810	1148
244	LI:816655.2:2000SEP08	6213342H1	811	1125
244	LI:816655.2:2000SEP08	6213374H1	811	1125
244	LI:816655.2:2000SEP08	g4290541	812	1287
244	LI:816655.2:2000SEP08	g6036605	812	1286
244	LI:816655.2:2000SEP08	1904405T6	773	1240
244	LI:816655.2:2000SEP08	g5363571	774	1285
244	LI:816655.2:2000SEP08	2252513H1	781	1081
244	LI:816655.2:2000SEP08	1362533H1	759	1034
244	LI:816655.2:2000SEP08	g2964362	760	1290
244	LI:816655.2:2000SEP08	g5364789	830	1285
244	LI:816655.2:2000SEP08	1602631F6	1	505
244	LI:816655.2:2000SEP08	g5589752	1056	1288

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
244	LI:816655.2:2000SEP08	g5233264	1000	1285
244	LI:816655.2:2000SEP08	556245H1	1044	1141
244	LI:816655.2:2000SEP08	041536H1	1050	1214
244	LI:816655.2:2000SEP08	g6197824	1007	1289
244	LI:816655.2:2000SEP08	g6086883	1011	1288
244	LI:816655.2:2000SEP08	5165489H2	1013	1153
244	LI:816655.2:2000SEP08	g4890484	1012	1285
244	LI:816655.2:2000SEP08	3801089H1	1014	1198
244	LI:816655.2:2000SEP08	g6074328	1067	1285
244	LI:816655.2:2000SEP08	735354H1	780	1026
244	LI:816655.2:2000SEP08	3712420H1	761	1088
244	LI:816655.2:2000SEP08	g3149389	767	1276
244	LI:816655.2:2000SEP08	5262589H2	769	1038
244	LI:816655.2:2000SEP08	g2555542	769	1285
244	LI:816655.2:2000SEP08	g3017238	763	1289
244	LI:816655.2:2000SEP08	g4409732	765	1285
244	LI:816655.2:2000SEP08	g2432338	765	1110
244	LI:816655.2:2000SEP08	1235291H1	766	1089
244	LI:816655.2:2000SEP08	g2716591	767	1289
244	LI:816655.2:2000SEP08	5668751T6	652	1231
244	LI:816655.2:2000SEP08	70444588D1	656	1280
244	LI:816655.2:2000SEP08	7384705H1	663	1244
244	LI:816655.2:2000SEP08	736319H1	780	1060
244	LI:816655.2:2000SEP08	1473108H1	782	1102
244	LI:816655.2:2000SEP08	406428H1	782	1028
244	LI:816655.2:2000SEP08	4704466H1	781	1078
244	LI:816655.2:2000SEP08	2856031H1	132	406
244	LI:816655.2:2000SEP08	3170550H1	427	633
245	LG:414732.1:2000SEP08	2525961F6	1	449
245	LG:414732.1:2000SEP08	2525961H1	2	243
245	LG:414732.1:2000SEP08	6560586H1	64	613
245	LG:414732.1:2000SEP08	g1357976	235	523
245	LG:414732.1:2000SEP08	2525961T6	253	677
245	LG:414732.1:2000SEP08	g3245391	532	894
246	LG:1140250.1:2000SEP08	g5438764	334	773
246	LG:1140250.1:2000SEP08	7453292H1	1	649
247	LG:174022.1:2000SEP08	6796523F8	1	606
247	LG:174022.1:2000SEP08	6796523H1	1	434
247	LG:174022.1:2000SEP08	5334862F6	226	806
247	LG:174022.1:2000SEP08	6797207H1	298	836
247	LG:174022.1:2000SEP08	6797207F8	301	728
247	LG:174022.1:2000SEP08	6797207T8	487	995
247	LG:174022.1:2000SEP08	5334862T6	576	1006
247	LG:174022.1:2000SEP08	g1557014	749	1188
247	LG:174022.1:2000SEP08	6796523T8	833	1016
248	LI:002811.1:2000SEP08	1444941H1	1	266
248	LI:002811.1:2000SEP08	4756409H1	16	214
248	LI:002811.1:2000SEP08	3865072F8	16	487
248	LI:002811.1:2000SEP08	6537148H1	16	309

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
248	LI:002811.1:2000SEP08	3088061F6	16	305
248	LI:002811.1:2000SEP08	3088061H1	16	228
248	LI:002811.1:2000SEP08	4756417H1	31	215
248	LI:002811.1:2000SEP08	5923138H1	84	359
248	LI:002811.1:2000SEP08	4314410H1	123	402
248	LI:002811.1:2000SEP08	4314410F8	123	645
248	LI:002811.1:2000SEP08	5994515F9	154	788
248	LI:002811.1:2000SEP08	5808417H1	154	320
248	LI:002811.1:2000SEP08	3865072T9	464	890
249	LI:414732.2:2000SEP08	70396476D1	483	676
249	LI:414732.2:2000SEP08	71650494V1	503	709
249	LI:414732.2:2000SEP08	g3245391	535	897
249	LI:414732.2:2000SEP08	70395826D1	279	680
249	LI:414732.2:2000SEP08	70397464D1	292	676
249	LI:414732.2:2000SEP08	70396886D1	292	564
249	LI:414732.2:2000SEP08	70395257D1	293	751
249	LI:414732.2:2000SEP08	70397538D1	293	768
249	LI:414732.2:2000SEP08	70394880D1	293	746
249	LI:414732.2:2000SEP08	70395475D1	298	676
249	LI:414732.2:2000SEP08	70395110D1	307	676
249	LI:414732.2:2000SEP08	70395982D1	312	680
249	LI:414732.2:2000SEP08	70397165D1	352	676
249	LI:414732.2:2000SEP08	71669221V1	419	624
249	LI:414732.2:2000SEP08	2525961H1	5	246
249	LI:414732.2:2000SEP08	70395557D1	40	642
249	LI:414732.2:2000SEP08	70397658D1	59	589
249	LI:414732.2:2000SEP08	6560586H1	67	616
249	LI:414732.2:2000SEP08	70396453D1	72	478
249	LI:414732.2:2000SEP08	70396167D1	119	654
249	LI:414732.2:2000SEP08	70396459D1	118	585
249	LI:414732.2:2000SEP08	70396121D1	192	481
249	LI:414732.2:2000SEP08	g1357976	238	526
249	LI:414732.2:2000SEP08	2525961T6	256	680
249	LI:414732.2:2000SEP08	70396053D1	260	679
249	LI:414732.2:2000SEP08	70396654D1	278	768
249	LI:414732.2:2000SEP08	70397446D1	279	676
249	LI:414732.2:2000SEP08	70395799D1	279	680
249	LI:414732.2:2000SEP08	70395867D1	1	542
249	LI:414732.2:2000SEP08	2525961F6	4	452
250	LI:1019920.1:2000SEP08	6797348T8	1	316
250	LI:1019920.1:2000SEP08	6797348F8	1	474
250	LI:1019920.1:2000SEP08	6797348H1	6	475
251	LI:1038336.1:2000SEP08	5759691F8	1	426
251	LI:1038336.1:2000SEP08	5759691T8	1	457
251	LI:1038336.1:2000SEP08	5759691H1	1	303
252	LI:1177772.11:2000SEP08	70784444V1	3476	3679
252	LI:1177772.11:2000SEP08	70784848V1	2920	3410
252	LI:1177772.11:2000SEP08	70782322V1	3075	3638
252	LI:1177772.11:2000SEP08	7195709H1	2072	2405

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
252	LI:1177772.11:2000SEP08	70646522V1	1543	1868
252	LI:1177772.11:2000SEP08	70783398V1	2920	3206
252	LI:1177772.11:2000SEP08	70784579V1	3553	3711
252	LI:1177772.11:2000SEP08	7286594H1	1543	2040
252	LI:1177772.11:2000SEP08	70542607V1	2426	2616
252	LI:1177772.11:2000SEP08	5609014H1	1	211
252	LI:1177772.11:2000SEP08	5609014F6	1	447
252	LI:1177772.11:2000SEP08	g1998690	31	378
252	LI:1177772.11:2000SEP08	g5511800	56	444
252	LI:1177772.11:2000SEP08	7468389H1	78	542
252	LI:1177772.11:2000SEP08	7243849H2	121	650
252	LI:1177772.11:2000SEP08	7243757H2	136	677
252	LI:1177772.11:2000SEP08	2136647H1	155	410
252	LI:1177772.11:2000SEP08	7042172H1	193	706
252	LI:1177772.11:2000SEP08	7604687J1	340	676
252	LI:1177772.11:2000SEP08	7004109H1	354	916
252	LI:1177772.11:2000SEP08	6894786J1	387	765
252	LI:1177772.11:2000SEP08	6455442H1	691	1218
252	LI:1177772.11:2000SEP08	6455442F8	726	1044
252	LI:1177772.11:2000SEP08	7468162H1	879	1360
252	LI:1177772.11:2000SEP08	2803772H1	968	1223
252	LI:1177772.11:2000SEP08	7405396H1	1169	1443
252	LI:1177772.11:2000SEP08	5643791R8	1309	1681
252	LI:1177772.11:2000SEP08	70781935V1	3170	3687
252	LI:1177772.11:2000SEP08	70544895V1	3589	3687
252	LI:1177772.11:2000SEP08	70545108V1	1667	1787
252	LI:1177772.11:2000SEP08	70646685V1	1735	2214
252	LI:1177772.11:2000SEP08	797593H1	1366	1492
252	LI:1177772.11:2000SEP08	70542501V1	1414	1874
252	LI:1177772.11:2000SEP08	6660222V1	1413	1868
252	LI:1177772.11:2000SEP08	4522328H1	1415	1658
252	LI:1177772.11:2000SEP08	4522328F6	1414	1703
252	LI:1177772.11:2000SEP08	6660385V1	1735	2228
252	LI:1177772.11:2000SEP08	70542862V1	1912	2597
252	LI:1177772.11:2000SEP08	70543554V1	1969	2573
252	LI:1177772.11:2000SEP08	70543721V1	2002	2581
252	LI:1177772.11:2000SEP08	70542966V1	2033	2732
252	LI:1177772.11:2000SEP08	70542802V1	2264	2754
252	LI:1177772.11:2000SEP08	5181114H2	2266	2474
252	LI:1177772.11:2000SEP08	70542604V1	2300	3004
252	LI:1177772.11:2000SEP08	355513H1	2309	2538
252	LI:1177772.11:2000SEP08	70542687V1	2426	3037
252	LI:1177772.11:2000SEP08	4736323H1	2444	2725
252	LI:1177772.11:2000SEP08	4609063H1	2492	2735
252	LI:1177772.11:2000SEP08	2079424H1	2517	2779
252	LI:1177772.11:2000SEP08	5913235H1	2589	2806
252	LI:1177772.11:2000SEP08	5848645H1	2600	2752
252	LI:1177772.11:2000SEP08	6271469H2	2624	3160
252	LI:1177772.11:2000SEP08	6271469F8	2631	3250

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
252	LI:1177772.11:2000SEP08	6060572H1	2716	3299
252	LI:1177772.11:2000SEP08	70542409V1	2780	3515
252	LI:1177772.11:2000SEP08	70782474V1	2920	3335
252	LI:1177772.11:2000SEP08	3278453F6	2920	3364
252	LI:1177772.11:2000SEP08	790881H1	2929	3179
252	LI:1177772.11:2000SEP08	790881R6	2929	3251
252	LI:1177772.11:2000SEP08	790881R1	2929	3461
252	LI:1177772.11:2000SEP08	5728562H1	2955	3424
252	LI:1177772.11:2000SEP08	673057H1	2972	3247
252	LI:1177772.11:2000SEP08	676375H1	2972	3238
252	LI:1177772.11:2000SEP08	2918392H1	3026	3307
252	LI:1177772.11:2000SEP08	6331008H1	3169	3681
252	LI:1177772.11:2000SEP08	6245944F8	3238	3804
252	LI:1177772.11:2000SEP08	6245944H1	3238	3742
252	LI:1177772.11:2000SEP08	6271469T8	3426	3876
252	LI:1177772.11:2000SEP08	4978711H1	3402	3670
252	LI:1177772.11:2000SEP08	6331008T8	3438	3886
252	LI:1177772.11:2000SEP08	3278453T6	3426	3748
252	LI:1177772.11:2000SEP08	790881F1	3465	3978
252	LI:1177772.11:2000SEP08	2821829H1	3461	3770
252	LI:1177772.11:2000SEP08	2821829F6	3461	3764
252	LI:1177772.11:2000SEP08	6245587T8	3482	3869
252	LI:1177772.11:2000SEP08	6553080T8	3492	3877
252	LI:1177772.11:2000SEP08	6245944T8	3495	3875
252	LI:1177772.11:2000SEP08	6940131H1	3500	3580
252	LI:1177772.11:2000SEP08	70542580V1	3540	3993
252	LI:1177772.11:2000SEP08	4522328T6	3556	3955
252	LI:1177772.11:2000SEP08	4832830F9	3566	3969
252	LI:1177772.11:2000SEP08	4832830F8	3566	3969
252	LI:1177772.11:2000SEP08	4832830H1	3566	3818
252	LI:1177772.11:2000SEP08	70544884V1	3589	3687
252	LI:1177772.11:2000SEP08	1370639H1	3611	3857
252	LI:1177772.11:2000SEP08	70791485V1	3601	3871
252	LI:1177772.11:2000SEP08	5643791F8	3621	3975
252	LI:1177772.11:2000SEP08	6331008F8	3731	3816
253	LI:205642.2:2000SEP08	6273658H1	1	560
253	LI:205642.2:2000SEP08	6274709F8	1	671
253	LI:205642.2:2000SEP08	6273658T8	71	710
254	LG:449685.1:2000SEP08	5912933T7	1	612
254	LG:449685.1:2000SEP08	5912933F7	13	613
254	LG:449685.1:2000SEP08	5912933H1	13	317
254	LG:449685.1:2000SEP08	5912933F8	13	462
254	LG:449685.1:2000SEP08	5912933T9	44	583
255	LG:453922.1:2000SEP08	5911278T9	74	634
255	LG:453922.1:2000SEP08	5911278T7	41	630
255	LG:453922.1:2000SEP08	5911278F7	1	591
255	LG:453922.1:2000SEP08	5911278T8	168	573
255	LG:453922.1:2000SEP08	5911278F8	1	569
255	LG:453922.1:2000SEP08	5911278H1	1	266

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
256	LG:476342.3:2000SEP08	5911370T8	1	438
256	LG:476342.3:2000SEP08	5911370F8	1	537
256	LG:476342.3:2000SEP08	5911370H1	1	286
256	LG:476342.3:2000SEP08	5911370F6	5	182
256	LG:476342.3:2000SEP08	6273147H1	9	534
256	LG:476342.3:2000SEP08	6271829F8	10	550
256	LG:476342.3:2000SEP08	6271829H1	10	534
256	LG:476342.3:2000SEP08	6271829T8	10	462
256	LG:476342.3:2000SEP08	6269367H1	14	496
256	LG:476342.3:2000SEP08	6268536H1	39	534
256	LG:476342.3:2000SEP08	6268536T8	39	431
256	LG:476342.3:2000SEP08	6268536F8	39	553
256	LG:476342.3:2000SEP08	5911370T6	98	495
257	LI:336801.1:2000SEP08	8032401J1	78	639
257	LI:336801.1:2000SEP08	g3048347	322	669
257	LI:336801.1:2000SEP08	g3049456	283	669
257	LI:336801.1:2000SEP08	g3240957	350	669
257	LI:336801.1:2000SEP08	g3921826	319	664
257	LI:336801.1:2000SEP08	g3888653	324	664
257	LI:336801.1:2000SEP08	g5768260	222	656
257	LI:336801.1:2000SEP08	g2115199	424	656
257	LI:336801.1:2000SEP08	g2458387	257	655
257	LI:336801.1:2000SEP08	g2458406	245	655
257	LI:336801.1:2000SEP08	g2882233	229	654
257	LI:336801.1:2000SEP08	g2907339	248	654
257	LI:336801.1:2000SEP08	g5850219	198	653
257	LI:336801.1:2000SEP08	g3149320	231	651
257	LI:336801.1:2000SEP08	g5765617	277	652
257	LI:336801.1:2000SEP08	g4648089	195	651
257	LI:336801.1:2000SEP08	g3753479	240	651
257	LI:336801.1:2000SEP08	g5393508	293	756
257	LI:336801.1:2000SEP08	g3037471	296	593
257	LI:336801.1:2000SEP08	g2115406	1	365
258	LI:449685.1:2000SEP08	5912933T7	1	616
258	LI:449685.1:2000SEP08	5912933F7	13	617
258	LI:449685.1:2000SEP08	5912933F8	13	464
258	LI:449685.1:2000SEP08	5912933H1	13	318
258	LI:449685.1:2000SEP08	5912933T9	44	587
259	LI:476342.1:2000SEP08	5321234F9	6	484
259	LI:476342.1:2000SEP08	5914061F8	6	444
259	LI:476342.1:2000SEP08	5913683H1	1	281
259	LI:476342.1:2000SEP08	5320619F9	1	372
259	LI:476342.1:2000SEP08	5913683F6	4	358
259	LI:476342.1:2000SEP08	5914061H1	6	264
259	LI:476342.1:2000SEP08	5913683T6	28	431
259	LI:476342.1:2000SEP08	5913683F8	27	200
259	LI:476342.1:2000SEP08	6269343F8	36	444
259	LI:476342.1:2000SEP08	6269343H1	36	444
259	LI:476342.1:2000SEP08	71637646V1	51	300

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
259	LI:476342.1:2000SEP08	6269670F8	59	444
259	LI:476342.1:2000SEP08	6269670T8	59	384
259	LI:476342.1:2000SEP08	6269670H1	59	444
259	LI:476342.1:2000SEP08	71605057V1	228	376
260	LI:1072804.1:2000SEP08	5910555F6	1	640
260	LI:1072804.1:2000SEP08	5910555F8	1	587
260	LI:1072804.1:2000SEP08	5910555T6	1	575
260	LI:1072804.1:2000SEP08	5910555T8	85	489
260	LI:1072804.1:2000SEP08	5910555T9	36	456
260	LI:1072804.1:2000SEP08	5910555H1	1	194
261	LI:455450.1:2000SEP08	5911845T6	1	432
261	LI:455450.1:2000SEP08	5911845F8	1	588
261	LI:455450.1:2000SEP08	5911845F6	1	484
261	LI:455450.1:2000SEP08	5911845H1	1	254
261	LI:455450.1:2000SEP08	5911845T8	22	443
262	LI:1073699.1:2000SEP08	6794742H1	22	365
262	LI:1073699.1:2000SEP08	6796418H1	22	365
262	LI:1073699.1:2000SEP08	6794009H1	22	297
262	LI:1073699.1:2000SEP08	6797985H1	59	383
262	LI:1073699.1:2000SEP08	6797985T8	59	257
262	LI:1073699.1:2000SEP08	6791482F8	59	365
262	LI:1073699.1:2000SEP08	6794247H1	1	362
262	LI:1073699.1:2000SEP08	6796383H1	22	365
262	LI:1073699.1:2000SEP08	6796341H1	59	366
262	LI:1073699.1:2000SEP08	6791482H1	59	365
262	LI:1073699.1:2000SEP08	6791482T8	59	255
262	LI:1073699.1:2000SEP08	6797985F8	59	390
262	LI:1073699.1:2000SEP08	6796341T8	59	249
262	LI:1073699.1:2000SEP08	6796341F8	59	365
263	LI:1013729.1:2000SEP08	8081248H2	1	644
263	LI:1013729.1:2000SEP08	6795338H1	49	567
263	LI:1013729.1:2000SEP08	6795338F8	49	643
263	LI:1013729.1:2000SEP08	6795338T8	376	468
264	LI:2050322.2:2000SEP08	1232566F1	464	1068
264	LI:2050322.2:2000SEP08	6430389H1	677	1254
264	LI:2050322.2:2000SEP08	1808715T6	691	1216
264	LI:2050322.2:2000SEP08	041912H1	694	986
264	LI:2050322.2:2000SEP08	5802291H1	646	971
264	LI:2050322.2:2000SEP08	6845563H1	795	1216
264	LI:2050322.2:2000SEP08	g3539276	814	1263
264	LI:2050322.2:2000SEP08	g5857161	818	1255
264	LI:2050322.2:2000SEP08	g7457026	823	1258
264	LI:2050322.2:2000SEP08	g4736196	823	1256
264	LI:2050322.2:2000SEP08	6019406H1	826	1259
264	LI:2050322.2:2000SEP08	g2883372	834	1257
264	LI:2050322.2:2000SEP08	955429H1	834	1147
264	LI:2050322.2:2000SEP08	5196594H1	835	1063
264	LI:2050322.2:2000SEP08	g6640997	837	1256
264	LI:2050322.2:2000SEP08	g5548433	837	1257



TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
264	LI:2050322.2:2000SEP08	g2896497	854	1256
264	LI:2050322.2:2000SEP08	1475376R6	473	961
264	LI:2050322.2:2000SEP08	4598585H1	473	767
264	LI:2050322.2:2000SEP08	1475376H1	473	674
264	LI:2050322.2:2000SEP08	2651228H1	502	759
264	LI:2050322.2:2000SEP08	4658940H1	537	768
264	LI:2050322.2:2000SEP08	4939484H1	543	780
264	LI:2050322.2:2000SEP08	2715435H1	558	813
264	LI:2050322.2:2000SEP08	1559421H1	599	821
264	LI:2050322.2:2000SEP08	6019901H1	603	1188
264	LI:2050322.2:2000SEP08	g652230	786	1241
264	LI:2050322.2:2000SEP08	2505229T6	672	1219
264	LI:2050322.2:2000SEP08	6430122H1	677	1272
264	LI:2050322.2:2000SEP08	3143329H1	779	1109
264	LI:2050322.2:2000SEP08	5754935T8	659	1139
264	LI:2050322.2:2000SEP08	4934014H1	661	961
264	LI:2050322.2:2000SEP08	1655116T6	666	1216
264	LI:2050322.2:2000SEP08	5525720H2	666	988
264	LI:2050322.2:2000SEP08	2734436H1	1091	1256
264	LI:2050322.2:2000SEP08	6586205H1	1	473
264	LI:2050322.2:2000SEP08	5291518F6	4	400
264	LI:2050322.2:2000SEP08	203885H1	289	541
264	LI:2050322.2:2000SEP08	5291518H1	4	254
264	LI:2050322.2:2000SEP08	71304357V1	291	961
264	LI:2050322.2:2000SEP08	70789164V1	303	535
264	LI:2050322.2:2000SEP08	g6577740	907	1256
264	LI:2050322.2:2000SEP08	g1406653	910	1259
264	LI:2050322.2:2000SEP08	5332444H1	919	1119
264	LI:2050322.2:2000SEP08	g3871471	921	1255
264	LI:2050322.2:2000SEP08	4737914H1	903	1097
264	LI:2050322.2:2000SEP08	g7319136	904	1254
264	LI:2050322.2:2000SEP08	g6036700	1059	1257
264	LI:2050322.2:2000SEP08	2358421H1	1061	1248
264	LI:2050322.2:2000SEP08	g4900806	1064	1248
264	LI:2050322.2:2000SEP08	g1885931	1076	1248
264	LI:2050322.2:2000SEP08	2358234H1	1078	1248
264	LI:2050322.2:2000SEP08	g3001843	1083	1259
264	LI:2050322.2:2000SEP08	2734436T6	1084	1218
264	LI:2050322.2:2000SEP08	4822254H1	1085	1352
264	LI:2050322.2:2000SEP08	2734436F6	1091	1256
264	LI:2050322.2:2000SEP08	3674083H1	469	689
264	LI:2050322.2:2000SEP08	g5865341	865	1256
264	LI:2050322.2:2000SEP08	g1425520	868	1259
264	LI:2050322.2:2000SEP08	g4004421	871	1255
264	LI:2050322.2:2000SEP08	g4266711	875	1264
264	LI:2050322.2:2000SEP08	g4003813	875	1255
264	LI:2050322.2:2000SEP08	g1953836	879	1255
264	LI:2050322.2:2000SEP08	g3539275	856	1256
264	LI:2050322.2:2000SEP08	6723335H1	858	1256

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
264	LI:2050322.2:2000SEP08	7020250H1	425	923
264	LI:2050322.2:2000SEP08	7934249H1	428	1097
264	LI:2050322.2:2000SEP08	7965961H1	447	1152
264	LI:2050322.2:2000SEP08	4185920H1	454	556
264	LI:2050322.2:2000SEP08	g6475407	1096	1256
264	LI:2050322.2:2000SEP08	3676421H1	469	761
264	LI:2050322.2:2000SEP08	70683609V1	465	1027
264	LI:2050322.2:2000SEP08	70683377V1	390	477
264	LI:2050322.2:2000SEP08	5393080H1	348	633
264	LI:2050322.2:2000SEP08	g2824231	967	1254
264	LI:2050322.2:2000SEP08	4185282H1	979	1214
264	LI:2050322.2:2000SEP08	2756901H1	985	1262
264	LI:2050322.2:2000SEP08	g2789017	1005	1256
264	LI:2050322.2:2000SEP08	3238912H1	1032	1256
264	LI:2050322.2:2000SEP08	g2752802	1046	1256
264	LI:2050322.2:2000SEP08	g4606954	941	1265
264	LI:2050322.2:2000SEP08	g1952200	947	1254
264	LI:2050322.2:2000SEP08	5392871H1	346	633
264	LI:2050322.2:2000SEP08	g5862585	729	1180
264	LI:2050322.2:2000SEP08	g5755440	739	1246
264	LI:2050322.2:2000SEP08	3859660H1	718	1016
264	LI:2050322.2:2000SEP08	2505229F6	630	1215
264	LI:2050322.2:2000SEP08	2505229H1	630	875
264	LI:2050322.2:2000SEP08	1811086T6	630	1227
264	LI:2050322.2:2000SEP08	g1527387	649	1101
264	LI:2050322.2:2000SEP08	2216534H1	629	887
264	LI:2050322.2:2000SEP08	4367794H1	624	899
264	LI:2050322.2:2000SEP08	2208322H1	629	896
264	LI:2050322.2:2000SEP08	1232566H1	464	707
264	LI:2050322.2:2000SEP08	1811086H1	617	878
264	LI:2050322.2:2000SEP08	2358421T6	702	1213
264	LI:2050322.2:2000SEP08	8056748J1	707	1257
264	LI:2050322.2:2000SEP08	70685857V1	606	797
264	LI:2050322.2:2000SEP08	1505277H1	615	892
264	LI:2050322.2:2000SEP08	4459805H1	616	897
264	LI:2050322.2:2000SEP08	1811086F6	617	1119
265	LI:891327.1:2000SEP08	4906137H2	1	289
265	LI:891327.1:2000SEP08	5427025T8	60	579
265	LI:891327.1:2000SEP08	5427025F8	61	677
265	LI:891327.1:2000SEP08	4906137F6	1	401
265	LI:891327.1:2000SEP08	5427025H1	61	343
265	LI:891327.1:2000SEP08	4516212H1	361	605
266	LI:2053076.1:2000SEP08	5617302H1	197	489
266	LI:2053076.1:2000SEP08	70536520V1	297	777
266	LI:2053076.1:2000SEP08	5617302R8	336	806
266	LI:2053076.1:2000SEP08	4317660T8	404	969
266	LI:2053076.1:2000SEP08	813140H1	451	710
266	LI:2053076.1:2000SEP08	4731735H1	649	916
266	LI:2053076.1:2000SEP08	5086779F8	1	563

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
266	LI:2053076.1:2000SEP08	7218403H1	551	992
266	LI:2053076.1:2000SEP08	4317660H1	252	508
266	LI:2053076.1:2000SEP08	g1880685	253	633
266	LI:2053076.1:2000SEP08	4179193H1	260	526
266	LI:2053076.1:2000SEP08	4935339H1	102	368
266	LI:2053076.1:2000SEP08	70535152V1	103	640
266	LI:2053076.1:2000SEP08	8037810H1	208	817
266	LI:2053076.1:2000SEP08	8037810J1	220	852
266	LI:2053076.1:2000SEP08	g1967076	225	580
266	LI:2053076.1:2000SEP08	4317660F8	255	695
266	LI:2053076.1:2000SEP08	7402448H1	1	297
266	LI:2053076.1:2000SEP08	6029531H1	2	317
266	LI:2053076.1:2000SEP08	5086779H1	1	246
266	LI:2053076.1:2000SEP08	6824284H1	49	510
266	LI:2053076.1:2000SEP08	4935339F6	103	396
267	LG:220085.1:2000SEP08	6541929H1	432	928
267	LG:220085.1:2000SEP08	g314066	261	662
267	LG:220085.1:2000SEP08	7083405H1	1	529
267	LG:220085.1:2000SEP08	g712155	128	380
267	LG:220085.1:2000SEP08	g712130	26	364
268	LG:406709.1:2000SEP08	g3738020	1024	1268
268	LG:406709.1:2000SEP08	g5596083	853	1268
268	LG:406709.1:2000SEP08	2007067H1	849	1063
268	LG:406709.1:2000SEP08	2010757H1	1038	1123
268	LG:406709.1:2000SEP08	6026715T8	609	1171
268	LG:406709.1:2000SEP08	g2035382	205	428
268	LG:406709.1:2000SEP08	6026715H1	1	224
268	LG:406709.1:2000SEP08	6026715F6	1	640
268	LG:406709.1:2000SEP08	g2115758	846	996
268	LG:406709.1:2000SEP08	7684247H1	129	721
268	LG:406709.1:2000SEP08	6026715F8	1	523
268	LG:406709.1:2000SEP08	6026715T6	719	1038
268	LG:406709.1:2000SEP08	g3052844	801	1268
268	LG:406709.1:2000SEP08	g5363052	814	1269
268	LG:406709.1:2000SEP08	2013065H1	1038	1288
268	LG:406709.1:2000SEP08	2013065T6	1038	1272
268	LG:406709.1:2000SEP08	g3232595	827	1272
268	LG:406709.1:2000SEP08	g4223124	1017	1270
268	LG:406709.1:2000SEP08	g2162155	894	1268
268	LG:406709.1:2000SEP08	2013065R6	1038	1323
268	LG:406709.1:2000SEP08	g5395578	797	1263
268	LG:406709.1:2000SEP08	g4487277	941	1265
269	LG:347863.9:2000SEP08	7966127H1	1	601
269	LG:347863.9:2000SEP08	5307074H1	227	442
269	LG:347863.9:2000SEP08	4947827H1	248	316
269	LG:347863.9:2000SEP08	4947827F8	266	745
269	LG:347863.9:2000SEP08	2936425H1	487	750
270	LI:1073027.1:2000SEP08	6792866H1	1	136
270	LI:1073027.1:2000SEP08	6792866F8	1	541

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
270	LI:1073027.1:2000SEP08	6797273T8	55	652
270	LI:1073027.1:2000SEP08	6792866T8	46	582
271	LI:347635.1:2000SEP08	5639879H1	258	492
271	LI:347635.1:2000SEP08	7757195H1	289	867
271	LI:347635.1:2000SEP08	6777521J1	435	1037
271	LI:347635.1:2000SEP08	g2743365	657	1116
271	LI:347635.1:2000SEP08	7152112H1	782	1203
271	LI:347635.1:2000SEP08	g2359789	785	1221
271	LI:347635.1:2000SEP08	7714779H1	877	1365
271	LI:347635.1:2000SEP08	g274093	891	1191
271	LI:347635.1:2000SEP08	5699927F6	1264	1717
271	LI:347635.1:2000SEP08	5699927H1	1265	1513
271	LI:347635.1:2000SEP08	7714779J1	223	794
271	LI:347635.1:2000SEP08	5639879F6	258	786
271	LI:347635.1:2000SEP08	7200294R8	1	617
271	LI:347635.1:2000SEP08	7757195J1	99	742
271	LI:347635.1:2000SEP08	7186104H1	201	742
272	LI:013685.1:2000SEP08	6912425J1	349	961
272	LI:013685.1:2000SEP08	3000601H1	668	974
272	LI:013685.1:2000SEP08	182766H1	860	1080
272	LI:013685.1:2000SEP08	7640576H1	871	1384
272	LI:013685.1:2000SEP08	182765H1	891	1090
272	LI:013685.1:2000SEP08	6912425F8	1204	1796
272	LI:013685.1:2000SEP08	6912425H1	1211	1796
272	LI:013685.1:2000SEP08	7323437H1	1588	2189
272	LI:013685.1:2000SEP08	7640576J1	1814	2218
272	LI:013685.1:2000SEP08	5631255H1	1866	2094
272	LI:013685.1:2000SEP08	5631387H1	1866	2071
272	LI:013685.1:2000SEP08	g6073415	1964	2157
272	LI:013685.1:2000SEP08	938941H1	371	654
272	LI:013685.1:2000SEP08	g2195009	438	746
272	LI:013685.1:2000SEP08	7248621H1	529	734
272	LI:013685.1:2000SEP08	6166114H1	39	340
272	LI:013685.1:2000SEP08	5951635F6	39	413
272	LI:013685.1:2000SEP08	g2197659	40	224
272	LI:013685.1:2000SEP08	5951635H1	39	157
272	LI:013685.1:2000SEP08	4332474H1	90	343
272	LI:013685.1:2000SEP08	4332474T6	333	798
272	LI:013685.1:2000SEP08	6912425R8	335	942
272	LI:013685.1:2000SEP08	6741026F8	1	231
272	LI:013685.1:2000SEP08	6622232H2	1	434
273	LI:406709.1:2000SEP08	2013065T6	1043	1279
273	LI:406709.1:2000SEP08	6026715T8	611	1178
273	LI:406709.1:2000SEP08	g4223124	1021	1277
273	LI:406709.1:2000SEP08	g3052844	804	1275
273	LI:406709.1:2000SEP08	g5363052	817	1276
273	LI:406709.1:2000SEP08	2010757H1	1043	1130
273	LI:406709.1:2000SEP08	2007067H1	852	1068
273	LI:406709.1:2000SEP08	g2115758	850	1000

TABLE 5

SEQ ID NO:	Template ID	Component ID	Start	Stop
273	LI:406709.1:2000SEP08	7684247H1	129	724
273	LI:406709.1:2000SEP08	6026715F8	1	525
273	LI:406709.1:2000SEP08	g2035382	205	430
273	LI:406709.1:2000SEP08	6026715H1	1	224
273	LI:406709.1:2000SEP08	2013065H1	1043	1295
273	LI:406709.1:2000SEP08	2013065R6	1043	1330
273	LI:406709.1:2000SEP08	g3232595	830	1279
273	LI:406709.1:2000SEP08	g2162155	898	1275
273	LI:406709.1:2000SEP08	g5596083	857	1275
273	LI:406709.1:2000SEP08	g3738020	1029	1275
273	LI:406709.1:2000SEP08	g4487277	945	1272
273	LI:406709.1:2000SEP08	g5395578	800	1270
274	LI:2052938.1:2000SEP08	71746272V1	150	745
274	LI:2052938.1:2000SEP08	6303292H1	125	424
274	LI:2052938.1:2000SEP08	71746949V1	150	747
274	LI:2052938.1:2000SEP08	71747212V1	150	742
274	LI:2052938.1:2000SEP08	71745617V1	149	731
274	LI:2052938.1:2000SEP08	g2955524	286	736
274	LI:2052938.1:2000SEP08	g4740512	304	734
274	LI:2052938.1:2000SEP08	71744444V1	150	574
274	LI:2052938.1:2000SEP08	71741423V1	150	553
274	LI:2052938.1:2000SEP08	2746019H1	261	490
274	LI:2052938.1:2000SEP08	g5231739	265	731
274	LI:2052938.1:2000SEP08	g6569347	274	734
274	LI:2052938.1:2000SEP08	5428676H1	150	401
274	LI:2052938.1:2000SEP08	71741551V1	171	225
274	LI:2052938.1:2000SEP08	71746335V1	152	592
274	LI:2052938.1:2000SEP08	55005970J1	247	829
274	LI:2052938.1:2000SEP08	2311091R6	409	732
274	LI:2052938.1:2000SEP08	2311091H1	409	653
274	LI:2052938.1:2000SEP08	71740956V1	419	719
274	LI:2052938.1:2000SEP08	633969H1	446	695
274	LI:2052938.1:2000SEP08	g6569255	472	734
274	LI:2052938.1:2000SEP08	5428676F6	150	641
274	LI:2052938.1:2000SEP08	6148913H1	1	463
274	LI:2052938.1:2000SEP08	7285289H1	1	437
275	LI:213208.1:2000SEP08	g5665419	1	440
275	LI:213208.1:2000SEP08	g5432007	2	419
275	LI:213208.1:2000SEP08	616667H1	13	132
275	LI:213208.1:2000SEP08	7677770H2	18	498
275	LI:213208.1:2000SEP08	g1364420	25	208
275	LI:213208.1:2000SEP08	g4734886	35	383
275	LI:213208.1:2000SEP08	3407953H1	74	340
275	LI:213208.1:2000SEP08	g2354840	83	383
275	LI:213208.1:2000SEP08	g4685246	129	383
275	LI:213208.1:2000SEP08	g3149184	242	383

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
1	LG:405741.3:2000SEP08	Digestive System - 31%, Urinary Tract - 25%
2	LG:337194.1:2000SEP08	Germ Cells - 48%, Unclassified/Mixed - 12%, Skin - 10%
3	LG:017108.4:2000SEP08	Sense Organs - 81%, Respiratory System - 13%
4	LG:372569.5:2000SEP08	Liver - 53%, Cardiovascular System - 16%
5	LG:968765.1:2000SEP08	Liver - 75%, Female Genitalia - 17%
6	LG:255999.16:2000SEP08	Nervous System - 100%
7	LG:977820.9:2000SEP08	Embryonic Structures - 22%, Pancreas - 22%, Musculoskeletal System - 15%
8	LI:1071608.1:2000SEP08	Liver - 77%, Pancreas - 13%
9	LI:1074023.1:2000SEP08	Liver - 100%
10	LI:453570.1:2000SEP08	Nervous System - 100%
11	LI:072072.1:2000SEP08	Hemic and Immune System - 17%, Germ Cells - 14%, Liver - 12%
12	LI:148565.4:2000SEP08	Nervous System - 57%, Male Genitalia - 43%
13	LI:368626.4:2000SEP08	Skin - 94%
14	LI:346123.1:2000SEP08	Exocrine Glands - 63%, Nervous System - 38%
15	LI:335795.11:2000SEP08	Female Genitalia - 20%, Nervous System - 15%
16	LI:246023.2:2000SEP08	Connective Tissue - 19%, Endocrine System - 18%
17	LG:1100661.1:2000SEP08	Liver - 100%
18	LG:475856.1:2000SEP08	Nervous System - 67%, Hemic and Immune System - 33%
19	LG:1015343.1:2000SEP08	Liver - 100%
20	LG:1400575.1:2000SEP08	Respiratory System - 29%, Male Genitalia - 24%, Endocrine System - 24%
21	LG:1080545.1:2000SEP08	Germ Cells - 35%, Urinary Tract - 15%, Liver - 12%, Pancreas - 12%
22	LG:213947.1:2000SEP08	Respiratory System - 50%, Hemic and Immune System - 50%
23	LI:720641.1:2000SEP08	Nervous System - 100%
24	LI:1023894.1:2000SEP08	Liver - 100%
25	LI:734904.1:2000SEP08	Unclassified/Mixed - 23%, Sense Organs - 20%
26	LI:1178118.1:2000SEP08	Sense Organs - 38%, Unclassified/Mixed - 17%, Endocrine System - 12%
27	LI:213947.1:2000SEP08	Respiratory System - 50%, Hemic and Immune System - 50%
28	LG:407304.1:2000SEP08	Endocrine System - 20%, Pancreas - 14%, Embryonic Structures - 14%
29	LG:337358.1:2000SEP08	Nervous System - 34%, Male Genitalia - 26%
30	LG:986090.1:2000SEP08	Nervous System - 100%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
31	LG:123250.1:2000SEP08	Musculoskeletal System - 32%, Hemic and Immune System - 32%, Respiratory System - 26%
32	LG:1028774.2:2000SEP08	Unclassified/Mixed - 23%, Female Genitalia - 13%, Male Genitalia - 12%
33	LG:338927.6:2000SEP08	Skin - 64%, Hemic and Immune System - 14%
34	LG:332944.2:2000SEP08	Endocrine System - 45%, Musculoskeletal System - 22%
35	LI:347174.5:2000SEP08	Skin - 39%, Male Genitalia - 14%
36	LI:477070.1:2000SEP08	Nervous System - 100%
37	LI:723144.1:2000SEP08	Nervous System - 100%
38	LI:1007188.1:2000SEP08	Liver - 100%
39	LI:1024412.1:2000SEP08	Liver - 100%
40	LI:284797.3:2000SEP08	Digestive System - 42%, Nervous System - 34%, Hemic and Immune System - 13%
41	LI:1092901.1:2000SEP08	Male Genitalia - 75%, Nervous System - 25%
42	LI:228930.1:2000SEP08	Nervous System - 55%, Nervous System - 18%, Respiratory System - 14%, Hemic and Immune System - 14%
43	LI:722913.1:2000SEP08	Nervous System - 100%
44	LG:457478.1:2000SEP08	Nervous System - 100%
45	LG:358719.1:2000SEP08	Urinary Tract - 96%
46	LG:105160.5:2000SEP08	Urinary Tract - 54%, Hemic and Immune System - 23%, Male Genitalia - 15%
47	LG:400705.1:2000SEP08	Endocrine System - 34%, Female Genitalia - 13%, Cardiovascular System - 11%, Urinary Tract - 11%, Hemic and Immune System - 11%
48	LG:221977.1:2000SEP08	Hemic and Immune System - 42%, Germ Cells - 16%
49	LG:898771.1:2000SEP08	Liver - 19%
50	LI:457478.1:2000SEP08	Liver - 34%, Pancreas - 31%, Nervous System - 17%
51	LI:125140.1:2000SEP08	Exocrine Glands - 81%
52	LI:021095.2:2000SEP08	Digestive System - 100%
53	LI:888730.1:2000SEP08	Liver - 81%, Endocrine System - 15%
54	LI:358719.1:2000SEP08	Urinary Tract - 96%
55	LI:351342.3:2000SEP08	Exocrine Glands - 48%, Female Genitalia - 15%
56	LI:256099.2:2000SEP08	Cardiovascular System - 27%, Germ Cells - 23%, Exocrine Glands - 15%
57	LI:2051991.1:2000SEP08	Hemic and Immune System - 47%, Urinary Tract - 21%, Digestive System - 16%, Respiratory System - 16%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
58	LG:980769.1:2000SEP08	Unclassified/Mixed - 68%, Male Genitalia - 11%, Hemic and Immune System - 11%
59	LG:332474.3:2000SEP08	Nervous System - 50%, Urinary Tract - 44%
60	LG:1087707.1:2000SEP08	Stomatognathic System - 74%
61	LG:415349.1:2000SEP08	Urinary Tract - 33%, Embryonic Structures - 21%, Unclassified/Mixed - 19%
62	LG:132420.2:2000SEP08	Respiratory System - 30%, Male Genitalia - 20%, Female Genitalia - 20%, Digestive System - 20%
63	LG:394201.1:2000SEP08	Embryonic Structures - 86%, Respiratory System - 14%
64	LG:1060884.1:2000SEP08	Germ Cells - 30%, Connective Tissue - 24%, Respiratory System - 15%
65	LG:242191.1:2000SEP08	Sense Organs - 44%, Endocrine System - 15%
66	LG:1063762.3:2000SEP08	Endocrine System - 23%, Embryonic Structures - 23%, Musculoskeletal System - 15%
67	LG:1100856.1:2000SEP08	Liver - 100%
68	LG:979390.2:2000SEP08	Liver - 26%, Pancreas - 26%, Connective Tissue - 20%
69	LG:1400447.1:2000SEP08	Respiratory System - 71%, Nervous System - 14%, Hemic and Immune System - 14%
70	LG:1400562.1:2000SEP08	Exocrine Glands - 19%, Respiratory System - 15%, Nervous System - 14%
71	LG:1076130.1:2000SEP08	Female Genitalia - 31%, Cardiovascular System - 25%, Exocrine Glands - 25%
72	LG:1064459.1:2000SEP08	Sense Organs - 62%, Endocrine System - 15%
73	LG:1079415.14:2000SEP08	Embryonic Structures - 90%, Nervous System - 10%
74	LG:1329431.3:2000SEP08	Respiratory System - 38%, Male Genitalia - 25%, Digestive System - 25%
75	LG:1088431.2:2000SEP08	Exocrine Glands - 50%, Cardiovascular System - 25%, Urinary Tract - 25%
76	LG:1329462.2:2000SEP08	Female Genitalia - 25%, Digestive System - 21%, Liver - 19%
77	LI:393468.1:2000SEP08	Unclassified/Mixed - 53%, Unclassified/Mixed - 31%, Urinary Tract - 12%
78	LI:722577.1:2000SEP08	Nervous System - 100%
79	LI:322783.16:2000SEP08	Connective Tissue - 88%
80	LI:901355.2:2000SEP08	Male Genitalia - 97%
81	LI:038859.2:2000SEP08	Unclassified/Mixed - 61%, Female Genitalia - 19%
82	LI:1046117.1:2000SEP08	Sense Organs - 85%, Germ Cells - 15%
83	LI:801015.1:2000SEP08	Male Genitalia - 100%
84	LI:1175590.1:2000SEP08	Musculoskeletal System - 57%, Endocrine System - 43%
85	LI:1170585.2:2000SEP08	Endocrine System - 38%, Endocrine System - 28%, Musculoskeletal System - 13%
86	LI:719531.2:2000SEP08	Hemic and Immune System - 100%
87	LI:794623.1:2000SEP08	Urinary Tract - 75%, Female Genitalia - 25%



TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
88	LI:1173119.1:2000SEP08	Digestive System - 32%, Liver - 26%, Respiratory System - 19%
89	LI:1093285.1:2000SEP08	Respiratory System - 36%, Female Genitalia - 32%, Digestive System - 28%
90	LI:1091881.1:2000SEP08	Hemic and Immune System - 45%, Female Genitalia - 27%, Respiratory System - 27%
91	LI:1091617.1:2000SEP08	Nervous System - 67%, Male Genitalia - 33%
92	LI:1082344.1:2000SEP08	Musculoskeletal System - 73%, Digestive System - 27%
93	LI:1166249.1:2000SEP08	Liver - 34%, Exocrine Glands - 29%, Endocrine System - 17%
94	LI:799675.1:2000SEP08	Female Genitalia - 70%, Endocrine System - 12%, Exocrine Glands - 10%
95	LI:1178899.1:2000SEP08	Female Genitalia - 69%, Female Genitalia - 16%
96	LI:1169241.1:2000SEP08	Unclassified/Mixed - 55%, Musculoskeletal System - 26%, Nervous System - 13%
97	LI:1180090.1:2000SEP08	Connective Tissue - 64%, Urinary Tract - 29%
98	LI:2049322.1:2000SEP08	Urinary Tract - 54%, Digestive System - 21%, Respiratory System - 13%, Male Genitalia - 13%
99	LI:809074.1:2000SEP08	Skin - 76%
100	LI:805158.1:2000SEP08	Exocrine Glands - 38%, Nervous System - 23%, Male Genitalia - 23%
101	LI:1172697.1:2000SEP08	Nervous System - 27%, Connective Tissue - 16%, Digestive System - 13%
102	LI:1174107.2:2000SEP08	Sense Organs - 69%, Unclassified/Mixed - 26%
103	LI:1177434.2:2000SEP08	Unclassified/Mixed - 42%, Female Genitalia - 27%, Embryonic Structures - 19%
104	LI:1184255.1:2000SEP08	Skin - 36%, Unclassified/Mixed - 34%, Connective Tissue - 18%
105	LI:1164555.1:2000SEP08	Female Genitalia - 100%
106	LI:238666.4:2000SEP08	Endocrine System - 29%, Embryonic Structures - 29%, Exocrine Glands - 12%
107	LI:1166752.1:2000SEP08	Endocrine System - 32%, Exocrine Glands - 26%, Nervous System - 21%, Urinary Tract - 21%
108	LI:2049654.1:2000SEP08	Respiratory System - 50%, Female Genitalia - 25%, Hemic and Immune System - 17%
109	LI:242665.2:2000SEP08	Endocrine System - 67%, Female Genitalia - 33%
110	LI:208637.1:2000SEP08	Cardiovascular System - 17%, Stomatognathic System - 16%, Liver - 11%
111	LI:2051808.1:2000SEP08	Liver - 100%
112	LI:1175136.1:2000SEP08	Cardiovascular System - 100%
113	LI:1177337.1:2000SEP08	Unclassified/Mixed - 40%, Nervous System - 30%, Respiratory System - 15%, Hemic and Immune System - 15%
114	LI:1165056.1:2000SEP08	Female Genitalia - 58%, Nervous System - 12%, Nervous System - 10%
115	LI:1175250.1:2000SEP08	Germ Cells - 83%, Digestive System - 10%
116	LI:1183192.1:2000SEP08	Cardiovascular System - 62%, Urinary Tract - 21%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
117	LI:1183325.1:2000SEP08	Female Genitalia - 38%, Connective Tissue - 20%, Male Genitalia - 20%
119	LI:813422.1:2000SEP08	Sense Organs - 44%, Connective Tissue - 33%
120	LI:1093049.6:2000SEP08	Germ Cells - 20%, Endocrine System - 15%, Unclassified/Mixed - 14%
121	LI:202192.4:2000SEP08	Female Genitalia - 92%
122	LG:1041854.1:2000SEP08	Liver - 82%, Unclassified/Mixed - 18%
123	LG:1100502.1:2000SEP08	Liver - 97%
124	LI:726414.1:2000SEP08	Nervous System - 100%
125	LI:400517.4:2000SEP08	Stomatognathic System - 58%, Embryonic Structures - 18%
126	LI:1078917.1:2000SEP08	Liver - 100%
127	LI:1012560.1:2000SEP08	Unclassified/Mixed - 40%, Nervous System - 30%, Male Genitalia - 15%, Digestive System - 15%
128	LI:427997.4:2000SEP08	Liver - 15%, Male Genitalia - 14%, Embryonic Structures - 11%
129	LI:197899.1:2000SEP08	Germ Cells - 24%, Male Genitalia - 19%, Unclassified/Mixed - 18%
130	LG:334199.1:2000SEP08	Unclassified/Mixed - 43%, Endocrine System - 17%, Liver - 12%
131	LG:334345.1:2000SEP08	Nervous System - 100%
132	LG:228092.1:2000SEP08	Liver - 33%, Unclassified/Mixed - 20%, Germ Cells - 16%
133	LG:098580.1:2000SEP08	Unclassified/Mixed - 59%, Cardiovascular System - 28%, Endocrine System - 14%
134	LG:969572.1:2000SEP08	Hemic and Immune System - 100%
135	LG:196958.1:2000SEP08	Hemic and Immune System - 53%, Germ Cells - 26%, Musculoskeletal System - 13%
136	LG:108781.1:2000SEP08	Hemic and Immune System - 21%, Connective Tissue - 11%
137	LG:1327885.1:2000SEP08	Liver - 100%
138	LI:449393.1:2000SEP08	Nervous System - 100%
139	LI:897616.1:2000SEP08	Liver - 98%
140	LI:736860.1:2000SEP08	Nervous System - 100%
141	LI:027066.6:2000SEP08	Respiratory System - 34%, Digestive System - 26%, Musculoskeletal System - 14%
142	LI:1074263.1:2000SEP08	Liver - 100%
143	LI:334345.1:2000SEP08	Nervous System - 100%
144	LI:1093914.1:2000SEP08	Nervous System - 52%, Unclassified/Mixed - 17%, Respiratory System - 13%
145	LI:1188168.1:2000SEP08	Exocrine Glands - 13%
146	LI:1065168.1:2000SEP08	Liver - 100%
147	LI:1180418.1:2000SEP08	Hemic and Immune System - 15%, Respiratory System - 13%, Musculoskeletal System - 13%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
148	LG:232648.1:2000SEP08	Exocrine Glands - 15%, Female Genitalia - 14%, Hemic and Immune System - 11%
149	LG:1078420.1:2000SEP08	Urinary Tract - 23%, Musculoskeletal System - 20%, Female Genitalia - 17%
150	LG:1397599.1:2000SEP08	Liver - 35%, Connective Tissue - 27%, Urinary Tract - 15%
151	LG:1397655.2:2000SEP08	Embryonic Structures - 53%, Female Genitalia - 19%, Urinary Tract - 11%
152	LG:241055.1:2000SEP08	Skin - 20%, Exocrine Glands - 17%, Hemic and Immune System - 13%, Endocrine System - 13%
153	LG:1101065.1:2000SEP08	Sense Organs - 42%, Unclassified/Mixed - 10%
154	LG:475629.1:2000SEP08	Nervous System - 100%
155	LI:348991.1:2000SEP08	Nervous System - 42%, Endocrine System - 25%, Male Genitalia - 21%
156	LI:475629.1:2000SEP08	Female Genitalia - 75%, Nervous System - 25%
157	LI:261331.1:2000SEP08	Hemic and Immune System - 100%
158	LI:815686.1:2000SEP08	Urinary Tract - 33%, Pancreas - 14%, Connective Tissue - 12%
159	LI:1167327.2:2000SEP08	Connective Tissue - 26%, Musculoskeletal System - 23%, Female Genitalia - 23%
160	LI:758009.3:2000SEP08	Respiratory System - 94%
161	LG:331593.1:2000SEP08	Hemic and Immune System - 39%, Unclassified/Mixed - 24%, Nervous System - 24%
162	LI:1094174.1:2000SEP08	Stomatognathic System - 30%, Musculoskeletal System - 11%
163	LI:814362.1:2000SEP08	Female Genitalia - 21%, Musculoskeletal System - 14%, Hemic and Immune System - 13%
164	LI:219542.1:2000SEP08	Unclassified/Mixed - 47%, Germ Cells - 31%, Male Genitalia - 20%
165	LI:726197.1:2000SEP08	Nervous System - 100%
166	LI:1075314.1:2000SEP08	Liver - 100%
167	LI:437883.1:2000SEP08	Liver - 98%
168	LG:336265.1:2000SEP08	Endocrine System - 19%, Musculoskeletal System - 18%, Embryonic Structures - 13%
169	LG:407788.2:2000SEP08	Embryonic Structures - 50%, Endocrine System - 22%, Male Genitalia - 11%, Digestive System - 11%
170	LG:1326925.1:2000SEP08	Liver - 47%, Digestive System - 42%, Male Genitalia - 11%
171	LI:332655.2:2000SEP08	Pancreas - 26%, Digestive System - 17%, Female Genitalia - 13%
172	LI:1184621.4:2000SEP08	Cardiovascular System - 71%, Endocrine System - 23%
173	LI:2051386.1:2000SEP08	Skin - 32%, Nervous System - 21%, Liver - 21%
174	LG:362757.1:2000SEP08	Connective Tissue - 78%, Nervous System - 22%
175	LG:406770.1:2000SEP08	Unclassified/Mixed - 29%, Urinary Tract - 21%, Female Genitalia - 16%
176	LG:1094640.1:2000SEP08	Musculoskeletal System - 87%, Digestive System - 13%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
177	LG:001929.1:2000SEP08	Stomatognathic System - 56%, Skin - 20%, Digestive System - 13%
178	LI:401322.1:2000SEP08	Sense Organs - 45%, Liver - 20%, Skin - 15%
179	LI:208748.1:2000SEP08	Unclassified/Mixed - 14%, Germ Cells - 13%, Connective Tissue - 10%
180	LI:407242.1:2000SEP08	Connective Tissue - 27%, Nervous System - 12%
181	LI:403409.1:2000SEP08	Stomatognathic System - 51%, Respiratory System - 10%
182	LI:450798.1:2000SEP08	Female Genitalia - 97%
183	LI:410317.1:2000SEP08	Skin - 64%, Hemic and Immune System - 13%
184	LI:340268.1:2000SEP08	Urinary Tract - 33%, Nervous System - 25%, Digestive System - 25%
185	LI:205167.1:2000SEP08	Pancreas - 18%, Respiratory System - 15%, Musculoskeletal System - 13%, Digestive System - 13%
186	LG:998844.1:2000SEP08	Germ Cells - 93%
187	LG:1043787.1:2000SEP08	Liver - 100%
188	LG:1098931.16:2000SEP08	Urinary Tract - 67%, Female Genitalia - 33%
189	LG:199423.2:2000SEP08	Hemic and Immune System - 100%
190	LI:1075297.1:2000SEP08	Hemic and Immune System - 100%
191	LI:1043321.1:2000SEP08	Liver - 100%
192	LI:297070.1:2000SEP08	Urinary Tract - 65%, Embryonic Structures - 15%
193	LI:1085041.1:2000SEP08	Liver - 100%
194	LI:1071544.1:2000SEP08	Liver - 100%
195	LI:2052480.1:2000SEP08	Digestive System - 39%, Digestive System - 21%, Liver - 17%
196	LG:450105.1:2000SEP08	Nervous System - 100%
197	LG:450581.1:2000SEP08	Nervous System - 100%
198	LG:450887.1:2000SEP08	Nervous System - 100%
199	LG:460809.1:2000SEP08	Exocrine Glands - 100%
200	LG:452089.1:2000SEP08	Nervous System - 100%
201	LG:1099416.1:2000SEP08	Embryonic Structures - 60%, Digestive System - 27%, Nervous System - 13%
202	LG:255713.1:2000SEP08	Male Genitalia - 44%, Respiratory System - 31%, Endocrine System - 25%
203	LG:998903.1:2000SEP08	Nervous System - 100%
204	LG:1119656.1:2000SEP08	Urinary Tract - 57%, Female Genitalia - 29%, Hemic and Immune System - 14%
205	LG:1096907.1:2000SEP08	Liver - 90%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
206	LG:1323741.1:2000SEP08	Liver - 52%, Connective Tissue - 32%
207	LG:1098372.1:2000SEP08	Nervous System - 50%, Hemic and Immune System - 50%
208	LG:1006783.1:2000SEP08	Liver - 100%
209	LG:1097562.1:2000SEP08	Liver - 100%
210	LG:998868.1:2000SEP08	Nervous System - 100%
211	LG:1063383.1:2000SEP08	Male Genitalia - 25%, Hemic and Immune System - 21%, Exocrine Glands - 14%, Urinary Tract - 14%
212	LG:1400567.1:2000SEP08	Digestive System - 57%, Female Genitalia - 29%, Hemic and Immune System - 14%
213	LI:449404.1:2000SEP08	Nervous System - 100%
214	LI:449941.2:2000SEP08	Nervous System - 100%
215	LI:450229.1:2000SEP08	Nervous System - 100%
216	LI:450399.3:2000SEP08	Nervous System - 100%
217	LI:455771.1:2000SEP08	Nervous System - 100%
218	LI:720459.1:2000SEP08	Endocrine System - 92%
219	LI:723156.1:2000SEP08	Nervous System - 100%
220	LI:728055.1:2000SEP08	Liver - 96%
221	LI:1020789.1:2000SEP08	Liver - 100%
222	LI:1071728.1:2000SEP08	Liver - 100%
223	LI:1084329.1:2000SEP08	Liver - 100%
224	LI:246422.1:2000SEP08	Hemic and Immune System - 67%, Nervous System - 33%
225	LI:1086066.1:2000SEP08	Liver - 100%
226	LI:223142.1:2000SEP08	Germ Cells - 51%, Female Genitalia - 24%
227	LI:885368.1:2000SEP08	Nervous System - 67%, Female Genitalia - 33%
228	LI:481782.1:2000SEP08	Cardiovascular System - 63%, Nervous System - 38%
229	LI:1093813.1:2000SEP08	Liver - 100%
230	LI:449413.2:2000SEP08	Nervous System - 100%
231	LI:450105.1:2000SEP08	Nervous System - 100%
232	LI:814285.1:2000SEP08	Liver - 71%, Male Genitalia - 29%
233	LI:1142855.1:2000SEP08	Digestive System - 83%, Nervous System - 17%
234	LI:817330.1:2000SEP08	Nervous System - 100%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
235	LI:817845.1:2000SEP08	Nervous System - 100%
236	LI:460809.1:2000SEP08	Exocrine Glands - 83%, Nervous System - 17%
237	LI:815874.1:2000SEP08	Musculoskeletal System - 90%
238	LI:255713.1:2000SEP08	Male Genitalia - 28%, Endocrine System - 21%, Respiratory System - 21%
239	LI:035973.1:2000SEP08	Digestive System - 60%, Embryonic Structures - 24%, Digestive System - 10%
240	LI:1138110.1:2000SEP08	Hemic and Immune System - 86%
241	LI:2049074.1:2000SEP08	Liver - 98%
242	LI:1092460.1:2000SEP08	Liver - 100%
243	LI:399421.1:2000SEP08	Unclassified/Mixed - 69%, Male Genitalia - 23%
244	LI:816655.2:2000SEP08	Female Genitalia - 11%
245	LG:414732.1:2000SEP08	Endocrine System - 82%, Nervous System - 18%
246	LG:1140250.1:2000SEP08	Respiratory System - 100%
247	LG:174022.1:2000SEP08	Sense Organs - 73%, Liver - 25%
248	LI:002811.1:2000SEP08	Nervous System - 33%, Endocrine System - 25%, Cardiovascular System - 21%, Female Genitalia - 21%
249	LI:414732.2:2000SEP08	Endocrine System - 80%, Nervous System - 20%
250	LI:101920.1:2000SEP08	Liver - 100%
251	LI:1038336.1:2000SEP08	Nervous System - 100%
252	LI:117772.11:2000SEP08	Male Genitalia - 37%, Female Genitalia - 22%
253	LI:205642.2:2000SEP08	Nervous System - 100%
254	LG:449685.1:2000SEP08	Nervous System - 100%
255	LG:453922.1:2000SEP08	Nervous System - 100%
256	LG:476342.3:2000SEP08	Nervous System - 100%
257	LI:336801.1:2000SEP08	Germ Cells - 70%, Unclassified/Mixed - 20%
258	LI:449685.1:2000SEP08	Nervous System - 100%
259	LI:476342.1:2000SEP08	Connective Tissue - 75%, Nervous System - 25%
260	LI:1072804.1:2000SEP08	Nervous System - 100%
261	LI:455450.1:2000SEP08	Nervous System - 100%
262	LI:1073699.1:2000SEP08	Liver - 100%
263	LI:1013729.1:2000SEP08	Liver - 73%, Endocrine System - 27%

TABLE 6

SEQ ID NO:	Template ID	Tissue Distribution
264	LI:2050322.2:2000SEP08	Pancreas - 35%
265	LI:891327.1:2000SEP08	Digestive System - 50%, Hemic and Immune System - 50%
266	LI:2053076.1:2000SEP08	Male Genitalia - 55%, Digestive System - 14%, Digestive System - 12%
267	LG:220085.1:2000SEP08	Hemic and Immune System - 57%, Digestive System - 29%, Nervous System - 14%
268	LG:406709.1:2000SEP08	Unclassified/Mixed - 66%, Male Genitalia - 29%
269	LG:347863.9:2000SEP08	Hemic and Immune System - 67%, Digestive System - 33%
270	LI:1073027.1:2000SEP08	Liver - 100%
271	LI:347635.1:2000SEP08	Female Genitalia - 49%, Musculoskeletal System - 22%
272	LI:013685.1:2000SEP08	Male Genitalia - 34%, Embryonic Structures - 28%, Endocrine System - 14%
273	LI:406709.1:2000SEP08	Unclassified/Mixed - 63%, Male Genitalia - 31%
274	LI:2052938.1:2000SEP08	Germ Cells - 63%, Endocrine System - 20%
275	LI:213208.1:2000SEP08	Germ Cells - 68%

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
276	3	133	3	401	g8996018	4.00E-21		hexokinase 1 isoform td
276	3	133	3	401	g8996017	4.00E-21		hexokinase 1 isoform ta/tb
276	3	133	3	401	g34670	4.00E-21		hexokinase type 1
277	1	160	250	729	g870752	2.00E-22		N-acetylglucosaminyltransferase V
277	1	160	250	729	g4545222	2.00E-22		alpha-1,3(6)-mannosylglycoprotein beta-1,6-N-acetylglucosaminyltransferase
277	1	160	250	729	g349091	5.00E-21		N-acetylglucosaminyltransferase V
278	2	125	239	613	g7687936	1.00E-18		possible adenylylate kinase
278	2	125	239	613	g10177920	3.00E-15		contains similarity to adenylylate kinase-gene_id:MCA23.18
278	2	125	239	613	g10176815	1.00E-14		adenylylate kinase-like
279	1	373	199	1317	g3273307	1.00E-148		Lysophospholipase
279	1	373	199	1317	g7290456	3.00E-82		CG6428 gene product
279	1	373	199	1317	g3874557	1.00E-81		(Z81041) predicted using GeneFinder~Similarity to E.coli L-asparaginase (SW:P18840), contains similarity to Pfam domain: PF00023 (Ank repeat), Score=65.5, E-value=3.7e-16, N=2; PF00710 (Asparaginase), Score=174.7, E-value=5.1e-49, N=1~cDNA EST yk9f7.3 comes from this gene~cDNA EST yk25c6.5 comes from this gene~cDNA EST yk128d6.3 comes from this gene~cDNA EST yk152f8.3 comes from this gene~cDNA EST yk348d9.3 comes from this gene~cDNA EST yk348d9.5 comes from this gene~cDNA EST yk225c12.3 comes from this gene~cDNA EST yk225c12.5 comes from this gene~cDNA EST yk430c7.5 comes from this gene
280	2	227	29	709	g488838	1.00E-105		CaBP1
280	2	227	29	709	g13905146	1.00E-105		Similar to protein disulfide isomerase-related protein
280	2	227	29	709	g12838858	1.00E-104		putative
282	2	281	17	859	g6996429	1.00E-111		dJ568C11.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthetase (acetate-coA ligase))
282	2	281	17	859	g12697774	1.00E-106		acetyl-CoA synthetase 2
282	2	281	17	859	g12697772	1.00E-104		acetyl-CoA synthetase 2



TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
283	3	115	198	542	g399660	3.00E-51		aldehyde reductase
283	3	115	198	542	g7677318	8.00E-51		aldehyde reductase
283	3	115	198	542	g12848322	8.00E-51		putative
284	3	214	30	671	g56336	1.00E-104		glutathione S-transferase (aa 1-209)
284	3	214	30	671	g459939	1.00E-104		glutathione S-transferase
284	3	214	30	671	g695303	2.00E-96		GST pi enzyme
285	3	212	3	638	g2909424	8.00E-86		Glyoxalase I
285	3	212	3	638	g12744892	5.00E-84		glyoxalase I
285	3	212	3	638	g2113825	3.00E-83		Glyoxalase I
286	2	183	44	592	g12856270	7.00E-06		putative
286	2	183	44	592	g10434969	7.00E-06		unnamed protein product
287	1	164	1	492	g414607	1.00E-09		glyceraldehyde-3-phosphate dehydrogenase
287	1	164	1	492	g409575	1.00E-09		glyceraldehyde-3-phosphate dehydrogenase
287	1	164	1	492	g312179	3.00E-09		glyceraldehyde 3-phosphate dehydrogenase
288	1	154	124	585	g13378170	7.00E-29		arachidonate lipooxygenase 3
288	1	154	124	585	g10799676	7.00E-29		lipooxygenase-3
288	1	154	124	585	g10441004	7.00E-29		epidermal lipooxygenase
289	2	148	2	445	g217974	1.00E-06		triosephosphate isomerase
289	2	148	2	445	g168647	1.00E-06		triosephosphate isomerase 1
290	2	291	485	1357	g12847081	1.00E-111		putative
290	2	291	485	1357	g12653491	1.00E-110		Similar to threonyl-tRNA synthetase
290	2	291	485	1357	g1464742	1.00E-109		threonyl-tRNA synthetase
291	1	354	1	1062	g10434528	0		unnamed protein product
291	1	354	1	1062	g13278319	1.00E-137		Similar to hypothetical protein FLJ12816
291	1	354	1	1062	g4929585	1.00E-111		CGI-58 protein
292	3	132	3	398	g57468	4.00E-63		oxytocin
292	3	132	3	398	g205900	4.00E-63		oxytocin/neurophysin
292	3	132	3	398	g205894	4.00E-63		oxytocin/neurophysin precursor
293	1	100	1	300	g6467206	2.00E-25		gonadotropin inducible transcription repressor-4
293	1	100	1	300	g6330394	8.00E-25		KIAA1198 protein

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
293	1	100	1	300	g12804721	3.00E-24	Unknown (protein for MGC:2663)
294	1	146	1	438	g57279	1.00E-51	pre-prosomatostatin
294	1	146	1	438	g297530	1.00E-51	pre-prosomatostatin
294	1	146	1	438	g207031	1.00E-51	somatostatin precursor
295	2	285	110	964	g6467206	1.00E-101	gonadotropin inducible transcription repressor-4
295	2	285	110	964	g13623354	4.00E-95	Similar to zinc finger protein 136 (clone pHZ-20)
295	2	285	110	964	g6330394	2.00E-93	KIAA1198 protein
296	2	213	2	640	g12052983	4.00E-71	hypothetical protein
296	2	213	2	640	g5262560	2.00E-68	hypothetical protein
296	2	213	2	640	g10434856	6.00E-68	unnamed protein product
297	1	95	94	378	g7262613	7.00E-13	candidate taste receptor T2R7
297	1	95	94	378	g7262619	2.00E-12	candidate taste receptor T2R10
297	1	95	94	378	g7262615	3.00E-12	candidate taste receptor T2R8
298	2	328	50	1033	g32093	8.00E-69	HGMP07J
298	2	328	50	1033	g1419016	7.00E-68	odorant receptor
298	2	328	50	1033	g11692549	2.00E-66	odorant receptor K30
299	3	101	660	962	g7158201	1.00E-25	cytokine receptor-like protein CYRL
301	3	427	348	1628	g14495650	2.00E-75	(BC009433) zinc finger protein 331; zinc finger protein 463
301	3	427	348	1628	g8575775	2.00E-75	KRAB zinc finger protein
301	3	427	348	1628	g13939858	2.00E-75	RITA
302	3	95	45	329	g7262613	7.00E-13	candidate taste receptor T2R7
302	3	95	45	329	g7262619	2.00E-12	candidate taste receptor T2R10
302	3	95	45	329	g7262615	3.00E-12	candidate taste receptor T2R8
304	2	407	221	1441	g9857402	1.00E-155	tumor endothelial marker 2
304	2	407	221	1441	g4092830	1.00E-155	dJ569D19.1 (similar to mouse Ras, Dexamethasone-induced 1 (Ras-related protein, RASD1, DEXRAS1))
304	2	407	221	1441	g5059122	1.00E-145	Rhes protein
305	1	277	1	831	g1519251	1.00E-131	GF14-c protein
305	1	277	1	831	g2921512	1.00E-123	GF14 protein
305	1	277	1	831	g7271253	1.00E-123	14-3-3-like protein

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
306	2	208	2	625	g6358507	7.00E-17		guanine exchange factor MCG7 isoform 2
306	2	208	2	625	g6358505	7.00E-17		guanine exchange factor MCG7 isoform 1
306	2	208	2	625	g4225848	7.00E-17		calcium- and diacylglycerol-regulated guanine nucleotide exchange factor 1
308	2	205	635	1249	g3599940	5.00E-78		faciogenital dysplasia protein 2
308	2	205	635	1249	g3342246	3.00E-50		actin-filament binding protein Frabin
308	2	205	635	1249	g3599944	1.00E-48		faciogenital dysplasia protein
309	1	545	1	1635	g5823454	5.00E-75		GTPase-activating protein 6 isoform 4
309	1	545	1	1635	g7243304	2.00E-74		rho-type GTPase-activating protein isoform 3
309	1	545	1	1635	g5881233	2.00E-74		rho GTPase-activating protein 6 isoform 5
310	2	212	2	637	g437985	1.00E-105		Rab12 protein
310	2	212	2	637	g206531	9.00E-96		RAB12
310	2	212	2	637	g12851149	4.00E-75		putative
311	2	189	2	568	g169931	3.00E-80		Glycine max calcium dependent protein kinase mRNA
311	2	189	2	568	g2501764	1.00E-75		calmodulin-like domain protein kinase isoenzyme beta
311	2	189	2	568	g7321076	5.00E-75		calmodulin-domain protein kinase CDPK isoform 4 (CPK4)
312	2	202	2	607	g2293566	1.00E-101		ADP-ribosylation factor 1
312	2	202	2	607	g2275195	1.00E-100		ADP-ribosylation factor 1
312	2	202	2	607	g166586	1.00E-100		ADP-ribosylation factor
313	1	178	1	534	g5689475	1.00E-67		KIAA1069 protein
313	1	178	1	534	g6705987	3.00E-48		phospholipase C-L2
313	1	178	1	534	g5689521	4.00E-48		KIAA1092 protein
314	3	77	75	305	g385168	3.00E-33		G-protein gamma subunit
314	3	77	75	305	g3450746	4.00E-31		GBG7_HUMAN
314	3	77	75	305	g3149954	4.00E-31		G-protein gamma 7
315	3	213	195	833	g13183338	1.00E-107		calneuron 1
315	3	213	195	833	g7670344	1.00E-107		unnamed protein product
315	3	213	195	833	g13183340	1.00E-107		calneuron 1
317	2	235	47	751	g9368450	1.00E-109		phospholipase C-beta-1b
317	2	235	47	751	g9368448	1.00E-109		phospholipase C-beta-1a

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
317	2	235	47	751	g206218	1.00E-108		phospholipase C-1
318	1	208	118	741	g3450893	1.00E-87		ras-like small monomeric GTP-binding protein
318	1	208	118	741	g7268592	4.00E-86		SAR1/GTP-binding secretory factor
318	1	208	118	741	g2104550	4.00E-86		AGAA.4
320	3	302	258	1163	g4584382	1.00E-152		E1a protein from 13s mma (32k, regulation and transformation)
320	3	302	258	1163	g209814	1.00E-149		32 kD protein
320	3	302	258	1163	g4584383	1.00E-115		E1a protein from 12s mma (26k, regulation and transformation)
321	1	162	1	486	g531901	2.00E-36		nuclear respiratory factor-2 subunit gamma 2
321	1	162	1	486	g531897	2.00E-36		nuclear respiratory factor-2 subunit beta 2
321	1	162	1	486	g286025	2.00E-36		E4TF1-53
322	1	195	34	618	g998899	5.00E-36		scleraxis=basic helix-loop-helix transcription factor (mice, embryos, Peptide, 207 aa)
322	1	195	34	618	g2155242	8.00E-22		paraxis
322	1	195	34	618	g1813563	2.00E-21		paraxis
324	1	262	1513	2298	g508528	1.00E-136		myocyte nuclear factor
324	1	262	1513	2298	g2289235	1.00E-119		myocyte nuclear factor-beta
324	1	262	1513	2298	g33854	2.00E-96		transcription factor ILF
326	3	408	165	1388	g6979924	1.00E-62		RP58
326	3	408	165	1388	g4959903	1.00E-62		transcriptional repressor RP58
326	3	408	165	1388	g4128145	1.00E-62		RP58 protein
327	1	59	1	177	g55624	2.00E-19		alpha initiation factor
327	1	59	1	177	g37058	2.00E-19		ILB protein
327	1	59	1	177	g339490	2.00E-19		transcription factor
329	3	104	144	455	g666914	8.00E-22		ferritin L-subunit
329	3	104	144	455	g309234	8.00E-22		ferritin light chain
329	3	104	144	455	g204133	8.00E-22		ferritin light chain
330	2	168	452	955	g4584382	2.00E-05		E1a protein from 13s mma (32k, regulation and transformation)
335	3	122	399	764	g4309888	2.00E-54		similar to zinc finger proteins; similar to protein S47071 (PID:g631503), match to EST AA339462 (NID:g1991774)
335	3	122	399	764	g9502403	2.00E-06		Hypothetical zinc finger-like protein

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
336	1	137	370	780	g186774	8.00E-27		zinc finger protein
336	1	137	370	780	g2384653	1.00E-25		Krueppel family zinc finger protein
336	1	137	370	780	g14348591	3.00E-25		KRAB zinc finger protein
338	2	166	2	499	g11345048	6.00E-26		SCAN domain-containing protein 2
338	2	166	2	499	g11320940	6.00E-26		SCAN2
338	2	166	2	499	g12859721	2.00E-25		putative
340	3	198	372	965	g7630121	1.00E-19		zinc finger protein 92
340	3	198	372	965	g1401082	1.00E-19		kruppel-type zinc finger protein
340	3	198	372	965	g2924250	3.00E-17		cJ29K1.2
341	2	439	341	1657	g3638956	0		zinc finger-like; similar to P52742 (PID:g1731411)
341	2	439	341	1657	g7670496	1.00E-170		unnamed protein product
341	2	439	341	1657	g1020145	7.00E-75		DNA binding protein
342	1	404	61	1272	g506502	1.00E-141		NK10
342	1	404	61	1272	g3135968	8.00E-71		b3418.1 (Kruppel related Zinc Finger protein 184)
342	1	404	61	1272	g3970712	2.00E-69		zinc finger protein 10
343	2	139	2	418	g13097465	6.00E-62		RIKEN cDNA 3110024A21 gene
343	2	139	2	418	g12837667	4.00E-61		putative
343	2	139	2	418	g2190184	3.00E-54		zinc finger protein
345	1	128	229	612	g9502403	2.00E-07		Hypothetical zinc finger-like protein
346	3	146	99	536	g2739353	9.00E-56		ZNF91L
346	3	146	99	536	g7959207	7.00E-50		KIAA1473 protein
346	3	146	99	536	g3342002	9.00E-50		hematopoietic cell derived zinc finger protein
347	3	250	3	752	g2843171	3.00E-71		zinc finger protein
347	3	250	3	752	g55471	6.00E-71		Zfp-29
347	3	250	3	752	g12855698	6.00E-71		putative
348	2	365	2	1096	g3135968	4.00E-84		b3418.1 (Kruppel related Zinc Finger protein 184)
348	2	365	2	1096	g5640017	7.00E-84		zinc finger protein ZFP113
348	2	365	2	1096	g1769491	2.00E-80		kruppel-related zinc finger protein
350	2	185	209	763	g4164083	1.00E-52		zinc finger protein EZNF
350	2	185	209	763	g2970038	1.00E-52		HKL1

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
350	2	185	209	763	g6007769	6.00E-42		KID1
351	1	89	127	393	g7023216	3.00E-18		unnamed protein product
351	1	89	127	393	g12804415	8.00E-18		Similar to hypothetical protein FLJ10891
351	1	89	127	393	g13752754	2.00E-16		zinc finger 1111
353	2	322	2	967	g2618752	1.00E-105		zinc finger protein
353	2	322	2	967	g6177785	4.00E-38		HKR1
353	2	322	2	967	g13325427	4.00E-38		Unknown (protein for IMAGE:3928207)
354	1	115	1	345	g5107180	1.00E-31		small zinc finger-like protein
354	1	115	1	345	g5107088	5.00E-30		small zinc finger-like protein
354	1	115	1	345	g5107174	3.00E-27		small zinc finger-like protein
355	1	108	379	702	g7159800	3.00E-58		dJ351K20.1.2 (novel C3HC4 type Zinc finger (RING finger) protein (isoform 2))
355	1	108	379	702	g7159799	3.00E-58		dJ351K20.1.1 (novel C3HC4 type Zinc finger (RING finger) protein (isoform 1))
355	1	108	379	702	g14198342	3.00E-58		hypothetical protein DKFZp434O1427
356	2	166	74	571	g8163824	1.00E-37		krueppel-like zinc finger protein HZF2
356	2	166	74	571	g498723	4.00E-32		zinc finger protein
356	2	166	74	571	g4235144	1.00E-28		BC39498_1
358	1	127	1	381	g13623633	1.00E-18		Unknown (protein for MGC:13105)
358	1	127	1	381	g4567180	1.00E-13		BC37295_2 (partial)
358	1	127	1	381	g12804721	6.00E-13		Unknown (protein for MGC:2663)
359	1	78	1	234	g487785	4.00E-16		zinc finger protein ZNF136
359	1	78	1	234	g13623607	4.00E-16		zinc finger protein 136 (clone pHZ-20)
359	1	78	1	234	g13623354	1.00E-15		Similar to zinc finger protein 136 (clone pHZ-20)
360	3	158	3	476	g14424716	2.00E-59		hypothetical protein FLJ11637
360	3	158	3	476	g10432938	2.00E-59		unnamed protein product
360	3	158	3	476	g9187356	2.00E-30		hypothetical protein, similar to (AB021644)GONADOTROPIN INDUCIBLE TRANSCRIPTION REPRESSOR-4
361	2	115	443	787	g7576274	2.00E-48		bA393J16.3 (novel KRAB box containing zinc finger gene)
361	2	115	443	787	g12841623	5.00E-34		putative

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
361	2	115	443	787	g488551	2.00E-25	zinc finger protein ZNF132
363	2	64	53	244	g7023216	1.00E-19	unnamed protein product
363	2	64	53	244	g12804415	1.00E-19	Similar to hypothetical protein FLJ10891
363	2	64	53	244	g13752754	7.00E-17	zinc finger 1111
364	3	255	3	767	g488551	2.00E-80	zinc finger protein ZNF132
364	3	255	3	767	g9968290	2.00E-78	zinc finger protein 304
364	3	255	3	767	g14249844	2.00E-78	Similar to hypothetical protein FLJ23233
365	1	97	253	543	g7023417	2.00E-36	unnamed protein product
365	1	97	253	543	g14042715	2.00E-36	unnamed protein product
365	1	97	253	543	g11917507	2.00E-36	HPF1 protein
366	1	158	1	474	g9802037	7.00E-61	zinc finger protein SBZF3
366	1	158	1	474	g4235144	1.00E-36	BC39498_1
366	1	158	1	474	g186774	3.00E-35	zinc finger protein
367	2	122	185	550	g5730196	2.00E-29	Kruppel-type zinc finger
367	2	122	185	550	g12849906	4.00E-28	putative
367	2	122	185	550	g55483	2.00E-24	Zfp-1 protein (AA 1-424)
368	1	242	1	726	g220637	1.00E-67	zinc finger protein
368	1	242	1	726	g4559318	2.00E-62	BC273239_1
368	1	242	1	726	g6467206	3.00E-62	gonadotropin inducible transcription repressor-4
369	3	92	48	323	g3342002	4.00E-37	hematopoietic cell derived zinc finger protein
369	3	92	48	323	g7959207	1.00E-35	KIAA1473 protein
369	3	92	48	323	g186774	2.00E-35	zinc finger protein
370	1	85	214	468	g7023216	4.00E-13	unnamed protein product
370	1	85	214	468	g12804415	4.00E-13	Similar to hypothetical protein FLJ10891
370	1	85	214	468	g13752754	6.00E-12	zinc finger 1111
372	3	206	3	620	g4567180	1.00E-102	BC37295_2 (partial)
372	3	206	3	620	g9502202	2.00E-99	endothelial zinc finger protein induced by tumor necrosis factor
372	3	206	3	620	g13879240	1.00E-16	Similar to zinc finger protein 46
373	1	206	121	738	g7981299	3.00E-41	dJ31316.6 (zinc finger protein 165)
373	1	206	121	738	g683471	3.00E-41	ZNF165

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
373	1	206	121	738	g4154166	3.00E-41	zinc finger protein
374	1	160	49	528	g13752754	1.00E-28	zinc finger 1111
374	1	160	49	528	g14348588	5.00E-28	KRAB zinc finger protein
374	1	160	49	528	g12654015	3.00E-27	Similar to hypothetical protein FLJ10891
375	2	115	140	484	g10434195	5.00E-57	unnamed protein product
375	2	115	140	484	g13529188	2.00E-36	Unknown (protein for MGC:12466)
375	2	115	140	484	g13623354	5.00E-33	Similar to zinc finger protein 136 (clone pHZ-20)
376	2	120	392	751	g14042186	3.00E-35	unnamed protein product
376	2	120	392	751	g55475	2.00E-32	Zink-finger protein 37
376	2	120	392	751	g53457	2.00E-32	zinc finger protein (AA 1-411)
377	1	273	73	891	g10435738	1.00E-142	unnamed protein product
377	1	273	73	891	g3342002	4.00E-87	hematopoietic cell derived zinc finger protein
377	1	273	73	891	g7959207	1.00E-85	KIAA1473 protein
378	3	132	48	443	g13623587	2.00E-45	Similar to zinc finger protein 254
378	3	132	48	443	g10435738	3.00E-35	unnamed protein product
378	3	132	48	443	g3342002	2.00E-34	hematopoietic cell derived zinc finger protein
379	2	233	563	1261	g2689444	5.00E-76	ZNF134
379	2	233	563	1261	g488553	7.00E-69	zinc finger protein ZNF134
379	2	233	563	1261	g10440218	3.00E-68	unnamed protein product
380	2	140	2	421	g7023417	2.00E-48	unnamed protein product
380	2	140	2	421	g14042715	2.00E-48	unnamed protein product
380	2	140	2	421	g11917507	2.00E-48	HPF1 protein
381	3	153	27	485	g10434781	9.00E-48	unnamed protein product
381	3	153	27	485	g13938351	5.00E-47	Similar to zinc finger protein 268
381	3	153	27	485	g7023216	3.00E-45	unnamed protein product
382	1	420	61	1320	g506502	1.00E-138	NK10
382	1	420	61	1320	g3135968	1.00E-66	b3418.1 (Kruppel related Zinc Finger protein 184)
382	1	420	61	1320	g1769491	8.00E-65	kruppel-related zinc finger protein
384	3	186	3	560	g9502403	3.00E-07	Hypothetical zinc finger-like protein



TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
386	1	476	43	1470	g2085786	0	similar to zinc finger 5 protein from Gallus gallus, U51640 (PID:g1399185)
386	1	476	43	1470	g4454855	0	zinc finger transcription factor Kaiso
386	1	476	43	1470	g1399187	2.00E-20	zinc finger 5 protein
387	1	292	1	876	g487284	1.00E-109	CRP2 (cysteine-rich protein 2)
387	1	292	1	876	g13938064	1.00E-108	RIKEN cDNA 0610010123 gene
387	1	292	1	876	g12832503	1.00E-108	putative
389	1	171	1	513	g186774	1.00E-71	zinc finger protein
389	1	171	1	513	g2723316	8.00E-71	Zinc-finger protein
389	1	171	1	513	g1017722	2.00E-69	repressor transcriptional factor
390	2	336	476	1483	g13752754	6.00E-72	zinc finger 1111
390	2	336	476	1483	g14348588	4.00E-69	KRAB zinc finger protein
390	2	336	476	1483	g10440398	4.00E-69	FLJ00032 protein
391	1	375	121	1245	g10434195	0	unnamed protein product
391	1	375	121	1245	g13529188	1.00E-128	Unknown (protein for MGC:12466)
391	1	375	121	1245	g6330394	1.00E-120	KIAA1198 protein
392	3	444	84	1415	g3135968	1.00E-102	b3418.1 (Kruppel related Zinc Finger protein 184)
392	3	444	84	1415	g14042550	1.00E-100	unnamed protein product
392	3	444	84	1415	g13937909	1.00E-100	Similar to KIAA0961 protein
393	1	620	235	2094	g14042330	0	unnamed protein product
393	1	620	235	2094	g3882317	0	KIAA0798 protein
393	1	620	235	2094	g12804011	0	KIAA0798 gene product
394	3	134	3	404	g487787	1.00E-13	zinc finger protein ZNF140
394	3	134	3	404	g13752754	5.00E-13	zinc finger 1111
394	3	134	3	404	g7023216	7.00E-13	unnamed protein product
396	2	114	413	754	g12853416	3.00E-24	putative
396	2	114	413	754	g13529497	6.00E-23	Unknown (protein for MGC:6652)
396	2	114	413	754	g4514561	1.00E-21	KRAB-containing zinc-finger protein KRAZ2
398	3	208	3	626	g457929	1.00E-65	delta subunit of F1F0 ATPase
398	3	208	3	626	g14198434	3.00E-63	RIKEN cDNA 0610008F14 gene

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
398	3	208	3	626	g12857538	3.00E-63		putative
399	3	104	51	362	g12859516	9.00E-49		putative
399	3	104	51	362	g12859507	9.00E-49		putative
399	3	104	51	362	g12859498	9.00E-49		putative
400	1	284	1	852	g3319340	3.00E-68		contains similarity to E. coli cation transport protein ChaC (GB:D90756)
400	1	284	1	852	g7270031	9.00E-67		predicted protein
400	1	284	1	852	g2827524	9.00E-67		predicted protein
401	2	297	2	892	g7290145	7.00E-46		EG:8D8.3 gene product
401	2	297	2	892	g2950398	7.00E-46		/prediction=(method:"gensecan", version:"1.0", score:"294.38")~/match=(desc:"THIAZIDE-SENSITIVE SODIUM-CHLORIDE COTRANSPORTER (NA-CL SYMPORTER)", species:"HOMO SAPIENS (HUMAN)", ranges:(query:33174..33518, target:SWISS-PROT::P55017:246..132, score:"258.00"), (query:33015..33149, target:SWISS-PROT::P55017:304..260, score:"75.00"), (query:32642..32761, target:SWISS-PROT::P55017:378..339, score:"121.00"), (query:32339..32503, target:SWISS-PROT::P55017:498..444, score:"76.00"), (query:32122..32268, target:SWISS-PROT::P55017:548..500, score:"97.00"), (query:31494..31625, target:SWISS-PROT::P55017:617..574, score:"55.00"), method:"blastx", version:"1.4.9")~/match=(desc:"BUMETANIDE-SENSITIVE SODIUM-(POTASSIUM)-CHLORIDE COTRANSPORTER 1 (BASOLATERAL NA-K-CL SYMPORTER)", species:"HOMO SAPIENS (HUMAN)", ranges:(query:33174..33512, target:SWISS-PROT::P55011:395..283, score:"247.00"), (query:33012..33149, target:SWISS-PROT::P55011:454..409, score:"97.00"), (query:32642..32827, target:SWISS-PROT::P55011:525..464, score:"154.00"), (query:32339..32515, target:SWISS-PROT::P55011:644..586, score:"93.00"), (query:32122..32268, target:SWISS-

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
401	2	297	2	892	g1086832	3.00E-37		coded for by <i>C. elegans</i> cDNA cm13g1; Similar to bumetanide-sensitive Na-K-Cl cotransporter
402	2	148	2	445	g510196	7.00E-84		TF
402	2	148	2	445	g1854476	7.00E-84		transferrin
402	2	148	2	445	g14250269	8.00E-72		Unknown (protein for IMAGE:3592890)
403	3	165	651	1145	g9957542	6.00E-86		connexin-59
403	3	165	651	1145	g14189950	6.00E-86		connexin 58
403	3	165	651	1145	g10946367	7.00E-36		connexin 55.5
404	3	285	3	857	g6996442	8.00E-49		CTL1 protein
404	3	285	3	857	g6996589	2.00E-47		CTL1 protein
404	3	285	3	857	g6996587	7.00E-39		CTL1 protein
405	2	414	749	1990	g14042129	0		unnamed protein product
405	2	414	749	1990	g2116552	1.00E-158		cationic amino acid transporter 3
405	2	414	749	1990	g1575776	1.00E-154		cationic amino acid transporter
407	2	188	2	565	g5921501	9.00E-21		distal intestinal serine protease
407	2	188	2	565	g4753837	8.00E-18		trypsin
407	2	188	2	565	g4753835	4.00E-17		trypsin
409	3	109	60	386	g999454	5.00E-30		TX protease precursor
409	3	109	60	386	g903934	5.00E-30		cysteine protease
409	3	109	60	386	g886050	5.00E-30		Ich-2
410	3	235	3	707	g205308	1.00E-140		alpha-1 major acute phase protein prepeptide
410	3	235	3	707	g207341	1.00E-133		T-kininogen
410	3	235	3	707	g205085	1.00E-131		LMW T-kininogen I precursor
411	2	210	11	640	g13516326	9.00E-20		marapsin
411	2	210	11	640	g12841953	1.00E-17		putative
411	2	210	11	640	g12836503	7.00E-16		putative
413	3	184	3	554	g4586674	2.00E-99		signal peptidase 21 kDa subunit
413	3	184	3	554	g12841311	2.00E-99		putative
413	3	184	3	554	g164084	2.00E-98		signal peptidase 21 kDa subunit
414	3	263	3	791	g6957716	1.00E-128		putative chaperonin

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
414	3	263	3	791	g14423532	1.00E-128		putative chaperonin
414	3	263	3	791	g9755653	1.00E-125		TCP-1 chaperonin-like protein
415	3	163	3	491	g433783	7.00E-76		binding protein
415	3	163	3	491	g337370	3.00E-74		rapamycin- and FK506-binding protein
415	3	163	3	491	g13097252	3.00E-74		Similar to FK506 binding protein 2 (13 kDa)
418	3	142	3	428	g736290	1.00E-66		precursor cystatin C C-terminal fragment (128 AA) (1 is 2nd base in codon)
418	3	142	3	428	g497415	2.00E-59		cystatin C
418	3	142	3	428	g12852172	2.00E-59		putative
419	1	191	88	660	g5921501	9.00E-21		distal intestinal serine protease
419	1	191	88	660	g4753837	8.00E-18		trypsin
419	1	191	88	660	g4753835	4.00E-17		trypsin
421	2	152	749	1204	g14043131	4.00E-20		Unknown (protein for IMAGE:2967328)
421	2	152	749	1204	g14017907	4.00E-20		KIAA1845 protein
421	2	152	749	1204	g13279050	4.00E-20		calpain 10
422	1	91	205	477	g56998	3.00E-33		proteasome subunit RC5
422	1	91	205	477	g3790135	6.00E-33		dj191N21.3 (proteasome subunit HC5)
422	1	91	205	477	g220026	6.00E-33		proteasome subunit C5
424	1	362	85	1170	g5748546	1.00E-159		C321D2.1 (Ribosomal Large Subunit Pseudouridine Synthase (EC 4.2.1.70, Pseudouridylylase Synthase, Uracil Hydrolase) LIKE protein)
424	1	362	85	1170	g14336724	1.00E-159		ribosomal large subunit pseudouridine synthase C like
424	1	362	85	1170	g12845023	1.00E-63		putative
425	2	169	119	625	g1263081	2.00E-77		mariner transposase
425	2	169	119	625	g3005702	1.00E-76		unknown
425	2	169	119	625	g2231380	9.00E-76		orf; encodes putative chimeric protein with SET domain in N-terminus with similarity to several other human, Drosophila, nematode and yeast proteins
428	1	131	184	576	g3005702	1.00E-44		unknown

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
428	1	131	184	576	g2231380	1.00E-44	orf; encodes putative chimeric protein with SET domain in N-terminus with similarity to several other human, Drosophila, nematode and yeast proteins
428	1	131	184	576	g1263081	1.00E-44	mariner transposase
430	2	101	74	376	g4185140	3.00E-40	putative small nuclear ribonucleoprotein E
430	2	101	74	376	g7269933	3.00E-38	small nuclear ribonucleoprotein homolog
430	2	101	74	376	g35105	3.00E-28	snRNP E protein (AA 1-92)
431	1	274	1	822	g3399667	1.00E-116	FBRL_HUMAN; 34 KD NUCLEOLAR SCLERODERMA ANTIGEN
431	1	274	1	822	g31395	1.00E-116	fibrillarin
431	1	274	1	822	g182592	1.00E-116	fibrillarin
432	3	110	24	353	g7269933	7.00E-26	small nuclear ribonucleoprotein homolog
432	3	110	24	353	g4185140	7.00E-26	putative small nuclear ribonucleoprotein E
432	3	110	24	353	g35105	8.00E-20	snRNP E protein (AA 1-92)
433	1	104	139	450	g2231380	4.00E-42	orf; encodes putative chimeric protein with SET domain in N-terminus with similarity to several other human, Drosophila, nematode and yeast proteins
433	1	104	139	450	g3005702	2.00E-41	unknown
433	1	104	139	450	g14286268	2.00E-41	SET domain and mariner transposase fusion gene
435	3	74	453	674	g2104910	4.00E-31	ORF derived from D1 leader region and Integrase coding region
435	3	74	453	674	g4959374	2.00E-21	pol protein
435	3	74	453	674	g2104914	2.00E-21	ORF derived from protease and integrase coding regions
436	2	187	293	853	g9650711	1.00E-72	HEF like Protein
436	2	187	293	853	g12964245	1.00E-72	dJ1167H4.4 (HEF like protein (HEFL))
436	2	187	293	853	g14042680	3.00E-67	unnamed protein product
437	3	187	147	707	g4106980	1.00E-06	Immunoglobulin-like transcript 10 protein
437	3	187	147	707	g3776468	2.00E-06	Immunoglobulin-like transcript 10 protein
437	3	187	147	707	g2645890	3.00E-06	IGSF1
439	3	175	3	527	g598166	5.00E-66	Immunoglobulin kappa chain variable region
439	3	175	3	527	g5360673	2.00E-64	anti-Entamoeba histolytica Immunoglobulin kappa light chain

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
439	3	175	3	527	g261240	5.00E-64	immunoglobulin M light chain V region=anti-lipid A antibody (human, hybridoma cell line HR78, Peptide Pair1d, 141 aa)
440	3	87	159	419	g30151	2.00E-26	cytochrome c oxidase subunit VIIb
440	3	87	159	419	g12834072	2.00E-25	putative
440	3	87	159	419	g12832690	2.00E-25	putative
441	3	119	3	359	g2114207	6.00E-49	glutaredoxin
441	3	119	3	359	g485953	4.00E-48	glutaredoxin
441	3	119	3	359	g10178147	2.00E-39	glutaredoxin-like protein
442	1	226	1	678	g1518874	2.00E-78	integral membrane protein CII-3
442	1	226	1	678	g13543226	3.00E-77	Similar to RIKEN cDNA 0610010E03 gene
442	1	226	1	678	g12849813	2.00E-76	putative
443	2	152	2	457	g205628	1.00E-27	24-kDa mitochondrial NADH dehydrogenase precursor (EC
443	2	152	2	457	g12850902	1.00E-27	putative
443	2	152	2	457	g3123721	2.00E-27	24-kDa subunit of complex I
445	3	169	903	1409	g5102636	3.00E-06	dJ682J15.1 (novel Collagen triple helix repeat containing protein)
445	3	169	903	1409	g12052774	3.00E-06	hypothetical protein
447	1	194	304	885	g12856559	1.00E-85	putative
447	1	194	304	885	g12856631	3.00E-85	putative
447	1	194	304	885	g12849896	3.00E-85	putative
450	1	160	1	480	g1419370	4.00E-74	actin depolymerizing factor
450	1	160	1	480	g10441256	5.00E-46	actin-depolymerizing factor 1
450	1	160	1	480	g9757910	2.00E-44	actin depolymerizing factor 4
452	3	267	3	803	g8249467	3.00E-31	titin
452	3	267	3	803	g1017427	7.00E-28	elastic titin
452	3	267	3	803	g9826	1.00E-07	11-1 polypeptide
453	1	525	37	1611	g52785	1.00E-109	57 kd keratin (aa 1-524)
453	1	525	37	1611	g386850	1.00E-109	keratin K5
453	1	525	37	1611	g34073	1.00E-109	cytokeratin 4 (408 AA)
455	1	154	1	462	g8777465	5.00E-58	cytoplasmic dynein heavy chain
455	1	154	1	462	g12852400	2.00E-33	putative

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
455	1	154	1	462	g3876099	4.00E-16		(Z75536) similar to dynein heavy chain-cDNA EST yk13d11.3 comes from this gene-cDNA EST yk13d11.5 comes from this gene
456	2	596	56	1843	g3769362	1.00E-70		ectoderm-neural cortex-1 protein
456	2	596	56	1843	g3309573	1.00E-70		nuclear matrix protein NRP/B
456	2	596	56	1843	g2282582	1.00E-70		actin-binding protein
457	3	762	3	2288	g8896164	0		kinesin-like protein GAKIN
457	3	762	3	2288	g10697238	0		KIF13A
457	3	762	3	2288	g12054032	0		KINESIN-13A2
458	3	255	3	767	g4415996	2.00E-36		beta-tubulin 4
458	3	255	3	767	g4098331	2.00E-36		beta-tubulin 5
458	3	255	3	767	g4098319	2.00E-36		beta-tubulin 1
459	1	156	211	678	g386847	2.00E-39		keratin
459	1	156	211	678	g34069	2.00E-39		keratin
459	1	156	211	678	g914833	3.00E-39		keratin type II
460	3	216	465	1112	g9864780	3.00E-14		beta-actin
460	3	216	465	1112	g4204812	4.00E-14		actin
460	3	216	465	1112	g8895873	5.00E-14		actin
461	3	160	3	482	g63805	7.00E-31		tensin
461	3	160	3	482	g619577	7.00E-31		cardiac muscle tensin
461	3	160	3	482	g212755	7.00E-31		tensin
462	3	97	3	293	g7259234	3.00E-22		contains transmembrane (TM) region
462	3	97	3	293	g12861877	3.00E-22		putative
462	3	97	3	293	g12837694	3.00E-22		putative
463	1	124	211	582	g7023973	1.00E-72		phospholipid hydroperoxide glutathione peroxidase
463	1	124	211	582	g406111	1.00E-72		phospholipid hydroperoxide glutathione peroxidase
463	1	124	211	582	g1063636	1.00E-72		phospholipid hydroperoxide glutathione peroxidase
464	1	68	25	228	g452316	6.00E-14		acetyl-CoA carboxylase
464	1	68	25	228	g2138330	6.00E-14		acetyl-CoA carboxylase
464	1	68	25	228	g1399290	6.00E-14		acetyl-CoA carboxylase beta
465	1	85	1	255	g5670328	2.00E-22		copine III

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
465	1	85	1	255	g3327086	2.00E-22		KIAA0636 protein
465	1	85	1	255	g6453711	8.00E-21		copine VII protein
466	2	172	2	517	g204491	8.00E-84		glutathione S-transferase
466	2	172	2	517	g5762309	3.00E-82		microsomal glutathione S-transferase
466	2	172	2	517	g12836829	8.00E-82		putative
468	1	110	604	933	g9650954	2.00E-26		beta-1,6-N-acetylglucosaminyltransferase B
468	1	110	604	933	g12860327	1.00E-12		putative
468	1	110	604	933	g9650956	3.00E-08		beta-1,6-N-acetylglucosaminyltransferase A
469	3	212	3	638	g206117	1.00E-97		prostaglandin H2 D-isomerase
469	3	212	3	638	g206115	1.00E-97		prostaglandin D synthetase
469	3	212	3	638	g895868	4.00E-86		prostaglandin D synthetase
470	2	178	2	535	g53354	1.00E-56		nucleoside diphosphate kinase B
470	2	178	2	535	g4467843	1.00E-56		NM23-H2 protein
470	2	178	2	535	g349476	1.00E-56		c-myc transcription factor
471	2	152	113	568	g163152	1.00E-13		hexokinase 1
471	2	152	113	568	g8996018	3.00E-13		hexokinase 1 isoform td
471	2	152	113	568	g8996017	3.00E-13		hexokinase 1 isoform ta/tb
472	2	164	2	493	g643074	5.00E-76		putative 40S ribosomal protein s12
472	2	164	2	493	g6716785	2.00E-75		40S ribosomal protein S23
472	2	164	2	493	g14532718	8.00E-75		(AY039983) unknown protein
474	3	122	3	368	g7629994	2.00E-34		60S RIBOSOMAL PROTEIN L36 homolog
474	3	122	3	368	g7413634	2.00E-33		60S ribosomal protein-like
474	3	122	3	368	g3236242	2.00E-33		60S ribosomal protein L36
475	3	101	3	305	g14586963	1.00E-22		(AF362574) M75
475	3	101	3	305	g57115	1.00E-22		ribosomal protein L31 (AA 1-125)
475	3	101	3	305	g36130	1.00E-22		ribosomal protein L31 (AA 1-125)
476	2	207	14	634	g7340874	1.00E-77		ESTs D15590(C0900), D48950(S15542), D22684(C0900) correspond to a region of the predicted gene. -Similar to Arabidopsis thaliana 60S ribosomal protein L11A (L16A). (P42795) (AY039570) AT4g18730/F28A21_140



TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
476	2	207	14	634	g9758681	6.00E-77	ribosomal protein L11-like
477	3	83	303	551	g12842823	2.00E-18	putative
477	3	83	303	551	g57121	4.00E-18	ribosomal protein L37
477	3	83	303	551	g461232	4.00E-18	ribosomal protein L37
478	1	75	256	480	g14586963	2.00E-09	(AF362574) M75
478	1	75	256	480	g57115	2.00E-09	ribosomal protein L31 (AA 1-125)
478	1	75	256	480	g36130	2.00E-09	ribosomal protein L31 (AA 1-125)
479	1	162	1	486	g5106775	1.00E-64	ribosomal protein S12
479	1	162	1	486	g4263712	3.00E-46	40S ribosomal protein S12
479	1	162	1	486	g6587799	1.00E-45	Strong similarity to gb   AF067732 ribosomal protein S12 from Hordeum vulgare. ESTs gb   T41772, gb   T42570, gb   A1999345, gb   T20784, gb   F20068 come from this gene.
481	1	108	1	324	g57123	7.00E-46	ribosomal protein L37a (AA 1-92)
481	1	108	1	324	g36134	7.00E-46	ribosomal protein L37a
481	1	108	1	324	g312414	7.00E-46	ribosomal protein L37a
482	2	142	68	493	g57702	3.00E-37	ribosomal protein L35 (AA 1-123)
482	2	142	68	493	g12849009	3.00E-37	putative
482	2	142	68	493	g12846227	3.00E-37	putative
483	1	82	70	315	g409074	4.00E-31	HBp15/L22
483	1	82	70	315	g409072	4.00E-31	HBp15/L22
483	1	82	70	315	g409070	4.00E-31	HBp15/L22
484	1	163	7	495	g12858199	3.00E-91	putative
484	1	163	7	495	g12842650	3.00E-91	putative
484	1	163	7	495	g12833292	3.00E-91	putative
485	2	118	110	463	g488415	9.00E-61	ribosomal protein L30
485	2	118	110	463	g3115336	9.00E-61	ribosomal protein L30
485	2	118	110	463	g206728	9.00E-61	ribosomal protein L30
486	2	260	2	781	g483431	1.00E-123	cyc07
486	2	260	2	781	g4079800	1.00E-122	S-phase-specific ribosomal protein
486	2	260	2	781	g6714564	1.00E-115	cyc07

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
487	1	115	640	984	g57117	1.00E-14		ribosomal protein L32
487	1	115	640	984	g36132	1.00E-14		rpl32 (aa 1-135)
487	1	115	640	984	g200781	1.00E-14		ribosomal protein L32-3A
488	2	57	125	295	g14586963	5.00E-19		(AF362574) M75
488	2	57	125	295	g57115	5.00E-19		ribosomal protein L31 (AA 1-125)
488	2	57	125	295	g36130	5.00E-19		ribosomal protein L31 (AA 1-125)
489	1	167	46	546	g4886269	9.00E-61		putative ribosomal protein S14
489	1	167	46	546	g12322890	3.00E-60		putative 40S ribosomal protein s14; 67401-66292
489	1	167	46	546	g4678226	1.00E-59		40S ribosomal protein S14
490	1	212	1	636	g1498053	3.00E-82		ribosomal protein S8
490	1	212	1	636	g968902	3.00E-73		ribosomal protein S8
490	1	212	1	636	g3264759	6.00E-73		40S ribosomal protein S8
491	1	217	1	651	g4588906	6.00E-97		ribosomal protein S7
491	1	217	1	651	g4128206	9.00E-83		40S ribosome protein S7
491	1	217	1	651	g3851636	9.00E-83		unknown
492	3	148	3	446	g13489168	5.00E-77		60S ribosomal protein L17
492	3	148	3	446	g13430182	3.00E-76		ribosomal protein L17
492	3	148	3	446	g14596111	2.00E-75		(AY042843) 60S ribosomal protein L17
493	3	158	3	476	g643074	5.00E-76		putative 40S ribosomal protein s12
493	3	158	3	476	g6716785	1.00E-75		40S ribosomal protein S23
493	3	158	3	476	g14532718	7.00E-75		(AY039983) unknown protein
494	1	188	1	564	g1490384	1.00E-104		ribosomal protein L6
494	1	188	1	564	g695638	1.00E-100		M-TAXREB107
494	1	188	1	564	g14210106	1.00E-100		ribosomal protein L6
495	3	144	3	434	g915313	2.00E-43		ribosomal protein L31
495	3	144	3	434	g7229709	3.00E-43		80S ribosomal protein L31
495	3	144	3	434	g2982295	3.00E-42		probable 60S ribosomal protein L31
496	3	159	3	479	g57714	5.00E-78		ribosomal protein S16 (AA 1-146)
496	3	159	3	479	g338447	5.00E-78		RPS16
496	3	159	3	479	g14044116	5.00E-78		ribosomal protein S16

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
497	1	147	1	441	g6006558	1.00E-74		ribosomal protein S18
497	1	147	1	441	g433447	1.00E-74		ribosomal protein S18
497	1	147	1	441	g4050105	1.00E-74		RPS18
498	1	130	58	447	g57720	8.00E-63		ribosomal protein S20 (AA 1-119)
498	1	130	58	447	g292443	8.00E-63		ribosomal protein S20
498	1	130	58	447	g13960133	8.00E-63		ribosomal protein S20
499	2	149	2	448	g57690	2.00E-67		ribosomal protein L23a
499	2	149	2	448	g404015	2.00E-67		ribosomal protein L23a
499	2	149	2	448	g306549	2.00E-67		homology to rat ribosomal protein L23
500	1	115	70	414	g409074	2.00E-40		HBp15/L22
500	1	115	70	414	g409072	2.00E-40		HBp15/L22
500	1	115	70	414	g409070	2.00E-40		HBp15/L22
501	2	178	2	535	g3717978	6.00E-86		5S ribosomal protein
501	2	178	2	535	g1685071	6.00E-86		ribosomal protein S5
501	2	178	2	535	g12861440	6.00E-86		putative
502	3	153	207	665	g57682	2.00E-18		ribosomal protein L17
502	3	153	207	665	g57111	2.00E-18		ribosomal protein L22
502	3	153	207	665	g34199	2.00E-18		putative ribosomal protein (AA 1-184)
503	2	155	2	466	g14596085	7.00E-67		(AY042830) Putative 40S ribosomal protein S15A
503	2	155	2	466	g9757906	7.00E-67		40S ribosomal protein S15A
503	2	155	2	466	g8439890	7.00E-67		Strong similarity to 40S ribosomal protein S15A from Arabidopsis thaliana gb L27461. EST gb R30315 comes from this gene.
504	2	98	272	565	g2331301	4.00E-41		ribosomal protein S4 type I
504	2	98	272	565	g2345154	3.00E-40		ribosomal protein S4
504	2	98	272	565	g457803	2.00E-38		ribosomal protein S4
505	1	132	1	396	g57702	3.00E-37		ribosomal protein L35 (AA 1-123)
505	1	132	1	396	g12849009	3.00E-37		putative
505	1	132	1	396	g12846227	3.00E-37		putative
506	3	163	12	500	g643074	5.00E-76		putative 40S ribosomal protein s12
506	3	163	12	500	g6716785	2.00E-75		40s ribosomal protein S23

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
506	3	163	12	500	g14532718	8.00E-75	(AY039983) unknown protein
507	1	119	139	495	g14532718	7.00E-33	(AY039983) unknown protein
507	1	119	139	495	g7413571	7.00E-33	putative protein
507	1	119	139	495	g6716785	7.00E-33	40s ribosomal protein S23
508	3	119	3	359	g14532718	7.00E-15	(AY039983) unknown protein
508	3	119	3	359	g7413571	7.00E-15	putative protein
508	3	119	3	359	g6716785	7.00E-15	40s ribosomal protein S23
509	1	107	334	654	g554269	2.00E-07	ribosomal protein L7
509	1	107	334	654	g36140	2.00E-07	ribosomal protein L7
509	1	107	334	654	g35903	2.00E-07	ribosomal protein L7
511	3	82	111	356	g5106775	7.00E-28	ribosomal protein S12
511	3	82	111	356	g6587799	6.00E-17	Strong similarity to gb AF067732 ribosomal protein S12 from Hordeum vulgare. ESTs gb T41772, gb T42570, gb AI999345, gb T20784, gb F20068 come from this gene.
511	3	82	111	356	g14190453	6.00E-17	At1g15930/T24D18_3
512	2	166	2	499	g9759463	2.00E-60	40S ribosomal protein S19
512	2	166	2	499	g6513924	4.00E-60	putative 40S ribosomal protein S19
512	2	166	2	499	g13878029	4.00E-60	putative 40S ribosomal protein S19
513	1	125	1	375	g14586963	2.00E-22	(AF362574) M75
513	1	125	1	375	g57115	2.00E-22	ribosomal protein L31 (AA 1-125)
513	1	125	1	375	g36130	2.00E-22	ribosomal protein L31 (AA 1-125)
514	3	91	216	488	g57720	2.00E-30	ribosomal protein S20 (AA 1-119)
514	3	91	216	488	g292443	2.00E-30	ribosomal protein S20
514	3	91	216	488	g214758	2.00E-30	ribosomal protein S22, 40S subunit
516	1	100	304	603	g12842823	3.00E-17	putative
516	1	100	304	603	g57121	7.00E-17	ribosomal protein L37
516	1	100	304	603	g461232	7.00E-17	ribosomal protein L37
518	2	188	2	565	g3869148	1.00E-10	ribosomal protein L13
518	2	188	2	565	g29383	1.00E-10	BBC1
518	2	188	2	565	g14043668	1.00E-10	ribosomal protein L13

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
519	1	143	1	429	G63466	2.00E-56		histone H2A
519	1	143	1	429	G6094631	2.00E-56		histone H2A.F
519	1	143	1	429	G3420799	2.00E-56		histone H2A.F/Z variant
520	3	197	951	1541	G14549638	2.00E-13		(AF255740) histone H2A variant
520	3	197	951	1541	G64777	2.00E-13		histone H2A (aa 1-130)
520	3	197	951	1541	G64325	2.00E-13		histone H2A
522	1	153	202	660	G183233	1.00E-33		beta-glucuronidase precursor (EC 3.2.1.31)
522	1	153	202	660	G14346709	1.00E-33		unnamed protein product
522	1	153	202	660	G3549609	2.00E-33		beta-glucuronidase
523	3	86	33	290	G8101071	3.00E-20		golgin-like protein
523	3	86	33	290	G8099669	3.00E-20		golgin-like protein
523	3	86	33	290	G7644350	2.00E-17		golgi matrix protein GM130
526	1	153	205	663	G183233	1.00E-33		beta-glucuronidase precursor (EC 3.2.1.31)
526	1	153	205	663	G14346709	1.00E-33		unnamed protein product
526	1	153	205	663	G3549609	2.00E-33		beta-glucuronidase
527	1	119	1	357	G4050095	2.00E-56		NADH oxidoreductase
527	1	119	1	357	G12845638	2.00E-56		putative
527	1	119	1	357	G12834155	2.00E-56		putative
528	2	62	35	220	G4454682	5.00E-07		NADH-ubiquinone oxidoreductase subunit B9 homolog
528	2	62	35	220	G4164444	5.00E-07		NADH:ubiquinone oxidoreductase B9 subunit
528	2	62	35	220	G248	1.00E-05		NADH dehydrogenase
529	1	381	2812	3954	G8101071	5.00E-18		golgin-like protein
529	1	381	2812	3954	G8099669	5.00E-18		golgin-like protein
529	1	381	2812	3954	G7644350	2.00E-17		golgi matrix protein GM130
530	3	173	3	521	G4510363	1.00E-67		putative ubiquitin-conjugating enzyme
530	3	173	3	521	G14596117	8.00E-67		(AY042846) Unknown protein
530	3	173	3	521	G4886271	8.00E-67		putative DNA-binding protein
531	3	124	3	374	G7248411	3.00E-38		ESTs C99632(E20954), C99633(E20954) correspond to a region of the predicted gene. ~Similar to Arabidopsis thaliana putative pathogenesis-related protein (U20347)

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability Score	Annotation
531	3	124	3	374	g7248405	1.00E-30	ESTs AU082419(E61744), AU031498(E61744) correspond to a region of the predicted gene.-Similar to Arabidopsis thaliana putative pathogenesis-related protein (U20347)
531	3	124	3	374	g6715639	4.00E-25	T25K16.16
532	1	135	88	492	g3789950	7.00E-56	translation initiation factor
532	1	135	88	492	g20238	7.00E-56	GOS2
532	1	135	88	492	g14194275	7.00E-56	translational initiation factor eIF1
535	1	172	82	597	g5912457	5.00E-87	dJ1068E13.2 (novel protein similar to bovine SCP2 (Sterol Carrier Protein 2) and part of HSD17B4 (hydroxysteroid (17-beta) dehydrogenase 4))
535	1	172	82	597	g12838636	1.00E-59	putative
535	1	172	82	597	g2315981	1.00E-35	17-beta-hydroxysteroid dehydrogenase type IV
538	2	168	2	505	g453189	1.00E-58	acyl carrier protein
538	2	168	2	505	g166971	4.00E-49	acyl carrier protein III
538	2	168	2	505	g166969	7.00E-41	acyl carrier protein II
541	3	146	3	440	g546420	1.00E-60	C-FABP=cutaneous fatty acid-binding protein (rats, Sprague-Dawley, skin, Peptide, 135 aa)
541	3	146	3	440	g1836058	4.00E-60	DA11=15.2 kDa fatty acid binding protein/FABP/C-FAPB homolog (rats, Sprague-Dawley, sciatic nerve traumatized, dorsal root ganglia, Peptide, 135 aa)
541	3	146	3	440	g533124	5.00E-60	lipid-binding protein
542	2	142	290	715	g7671659	1.00E-26	dJ1069P2.3.4 (novel PABPC1 (poly(A)-binding protein, cytoplasmic 1) (PABPL1) like protein (putative isoform 4))
542	2	142	290	715	g7671658	1.00E-26	dJ1069P2.3.3 (novel PABPC1 (poly(A)-binding protein, cytoplasmic 1) (PABPL1) like protein (putative isoform 3))
542	2	142	290	715	g7671657	1.00E-26	dJ1069P2.3.2 (novel PABPC1 (poly(A)-binding protein, cytoplasmic 1) (PABPL1) like protein (putative isoform 2))
543	3	175	3	527	g8176525	2.00E-23	Interferon-inducible myeloid differentiation transcriptional IFI16b
543	3	175	3	527	g6644297	2.00E-23	Interferon-gamma induced protein
543	3	175	3	527	g184569	2.00E-23	

TABLE 7

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability	Score	Annotation
544	3	96	639	926	g3879684	3.00E-05		(Z74D42) predicted using Genefinder~Similarity to Haemophilus 3-oxoacyl-(acyl-carrier protein) reductase (SW:FABG_HAEIN), contains similarity to Pfam domain: PF00106 (short chain dehydrogenase), Score=170.5, E-value=9.2e-48, N=1 ~cDNA EST yk470b2.3 comes from this gene~cDNA EST yk470b2.5 comes from (Human endogenous retrovirus type C oncovirus sequence.), gene product
545	3	139	219	635	g325465	9.00E-45		protease/polymerase
545	3	139	219	635	g2393895	2.00E-21		gag protein
545	3	139	219	635	g334989	8.00E-20		putative
546	3	219	3	659	g12855728	4.00E-72		putative
546	3	219	3	659	g12839570	4.00E-72		(AJ312322) OVARIAN/Breast septin beta
546	3	219	3	659	g14530111	7.00E-50		fornin-like protein
547	3	161	42	524	g6708478	1.00E-22		lymphocyte specific fomin related protein
547	3	161	42	524	g4101720	1.00E-21		CG6807 gene product
547	3	161	42	524	g7294416	2.00E-09		fertility protein SP22
548	1	177	1	531	g5478757	1.00E-91		fertility protein SP22
548	1	177	1	531	g5478755	1.00E-91		CAPI
548	1	177	1	531	g3250916	1.00E-91		NOTCH2 protein
549	2	241	2	724	g11527997	1.00E-150		NOTCH 2
549	2	241	2	724	g11275978	1.00E-150		Notch B
549	2	241	2	724	g287990	1.00E-149		putative
551	3	236	3	710	g12855728	1.00E-25		putative
551	3	236	3	710	g12839570	1.00E-25		(AJ312322) OVARIAN/Breast septin beta
551	3	236	3	710	g14530111	7.00E-14		differentiation enhancing factor 1
552	1	137	217	627	g4406393	8.00E-47		ADP-ribosylation factor-directed GTPase activating protein
552	1	137	217	627	g4063616	8.00E-47		ADP-ribosylation factor-directed GTPase activating protein
552	1	137	217	627	g4063614	8.00E-47		unnamed protein product
553	3	150	48	497	g14042295	2.00E-30		MSP028
553	3	150	48	497	g11640564	2.00E-30		Similar to tumor necrosis factor, alpha-induced protein 1
553	3	150	48	497	g13905272	3.00E-30		

Table 8

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	ESTs: fasta E value=1.0E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) <i>Nucleic Acids Res.</i> 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Attwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) <i>Nucleic Acids Res.</i> 26:320-322; Durbin, R. et al. (1998) <i>Our World View, in a Nutshell</i> , Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater



Table 8 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score > GCG-specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score = 3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

## CLAIMS

## What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:
  - 5 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275,
  - b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-275,
  - 10 c) a polynucleotide complementary to the polynucleotide of a),
  - d) a polynucleotide complementary to the polynucleotide of b), and
  - e) an RNA equivalent of a)-d).
2. An isolated polynucleotide of claim 1, selected from the group consisting of SEQ ID  
15 NO:1-275.
3. An isolated polynucleotide comprising at least 30 contiguous nucleotides of a polynucleotide of claim 1.
- 20 4. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 1.
5. A composition for the detection of expression of diagnostic and therapeutic polynucleotides comprising at least one of the polynucleotides of claim 1 and a detectable label.  
25
6. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 1, the method comprising:
  - a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
  - 30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
7. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 1, the method comprising:
  - 35 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and

which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

8. A method of claim 7, wherein the probe comprises at least 30 contiguous nucleotides.

9. A method of claim 7, wherein the probe comprises at least 60 contiguous nucleotides.

10. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 1.

11. A cell transformed with a recombinant polynucleotide of claim 10.

12. A transgenic organism comprising a recombinant polynucleotide of claim 10.

13. A method for producing a diagnostic and therapeutic polypeptide encoded by a polynucleotide of claim 1, the method comprising:

- a) culturing a cell under conditions suitable for expression of the diagnostic and therapeutic polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 1, and  
b) recovering the diagnostic and therapeutic polypeptide so expressed.

14. A method of claim 13, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

15. An isolated diagnostic and therapeutic polypeptide (DITHP) encoded by at least one of the polynucleotides of claim 2.

16. A method of screening for a test compound that specifically binds to the polypeptide of claim 15, the method comprising:

- a) combining the polypeptide of claim 15 with at least one test compound under suitable conditions, and  
b) detecting binding of the polypeptide of claim 15 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 15.

17. A microarray wherein at least one element of the microarray is a polynucleotide of claim 3.

5 18. A method for generating a transcript image of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- b) contacting the elements of the microarray of claim 17 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex,  
10 and
- c) quantifying the expression of the polynucleotides in the sample.

19. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence of a  
15 polynucleotide of claim 1, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying  
20 amounts of the compound and in the absence of the compound.

20. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at  
25 least 20 contiguous nucleotides of a polynucleotide of claim 1 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 1 or fragment thereof,
- c) quantifying the amount of hybridization complex, and
- d) comparing the amount of hybridization complex in the treated biological sample with  
30 the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

35 21. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first

oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 1.

5           22. An array of claim 21, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

          23. An array of claim 21, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide

10

          24. An array of claim 21, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

          25. An array of claim 21, which is a microarray.

15

          26. An array of claim 21, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

          27. An array of claim 21, wherein a linker joins at least one of said nucleotide molecules to  
20 said solid substrate.

          28. An array of claim 21, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains  
25 nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

          29. An isolated polypeptide selected from the group consisting of:

- 30           a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:276-553,
- b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:276-553,
- c) a biologically active fragment of a polypeptide having an amino acid sequence selected  
35 from the group consisting of SEQ ID NO:276-553, and
- d) an immunogenic fragment of a polypeptide having an amino acid sequence selected

from the group consisting of SEQ ID NO:276-553.

30. An isolated polypeptide of claim 29, having a sequence selected from the group consisting of SEQ ID NO:276-553.

5

31. An isolated polynucleotide encoding a polypeptide of claim 29.

32. An isolated polynucleotide encoding a polypeptide of claim 30.

10 33. An isolated polynucleotide of claim 32, having a sequence selected from the group consisting of SEQ ID NO:1-275.

34. An isolated antibody which specifically binds to a diagnostic and therapeutic polypeptide of claim 29.

15

35. A diagnostic test for a condition or disease associated with the expression of DITHP in a biological sample, the method comprising:

- 20 a) combining the biological sample with an antibody of claim 34, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

36. The antibody of claim 34, wherein the antibody is:

25

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a F(ab')<sub>2</sub> fragment, or
- e) a humanized antibody.

30

37. A composition comprising an antibody of claim 34 and an acceptable excipient.

38. A method of diagnosing a condition or disease associated with the expression of DITHP in a subject, comprising administering to said subject an effective amount of the composition of claim

35

37.

39. A composition of claim 37, wherein the antibody is labeled.

40. A method of diagnosing a condition or disease associated with the expression of DITHP in a subject, comprising administering to said subject an effective amount of the composition of claim 39.

41. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 34, the method comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

42. An antibody produced by a method of claim 41.

43. A composition comprising the antibody of claim 42 and a suitable carrier.

44. A method of making a monoclonal antibody with the specificity of the antibody of claim 34, the method comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

45. A monoclonal antibody produced by a method of claim 44.

46. A composition comprising the antibody of claim 45 and a suitable carrier.

47. The antibody of claim 34, wherein the antibody is produced by screening a Fab expression library.

48. The antibody of claim 34, wherein the antibody is produced by screening a recombinant immunoglobulin library.

49. A method of detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553 in a sample, the method comprising:

- a) incubating the antibody of claim 34 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553 in the sample.

50. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553 from a sample, the method comprising:

- a) incubating the antibody of claim 34 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:276-553.

51. A composition comprising a polypeptide of claim 29 and a pharmaceutically acceptable excipient.

52. A composition of claim 51, wherein the polypeptide has an amino acid sequence of SEQ ID NO:276-553.

53. A method for treating a disease or condition associated with decreased expression of functional DITHP, comprising administering to a patient in need of such treatment the composition of claim 51.

54. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 29, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 29 to a compound, and
- b) detecting agonist activity in the sample.



55. A composition comprising an agonist compound identified by a method of claim 54 and a pharmaceutically acceptable excipient.

56. A method for treating a disease or condition associated with decreased expression of functional DITHP, comprising administering to a patient in need of such treatment a composition of claim 55.

57. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 29, the method comprising:

- 10 a) exposing a sample comprising a polypeptide of claim 29 to a compound, and
- b) detecting antagonist activity in the sample.

58. A composition comprising an antagonist compound identified by a method of claim 57 and a pharmaceutically acceptable excipient.

59. A method for treating a disease or condition associated with overexpression of functional DITHP, comprising administering to a patient in need of such treatment a composition of claim 58.

60. A method of screening for a compound that modulates the activity of the polypeptide of claim 29, said method comprising:

- a) combining the polypeptide of claim 29 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 29,
- b) assessing the activity of the polypeptide of claim 29 in the presence of the test compound, and
- 25 c) comparing the activity of the polypeptide of claim 29 in the presence of the test compound with the activity of the polypeptide of claim 29 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 29 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 29.

<110> INCYTE GENOMICS, INC.  
JACKSON, Stuart  
LINCOLN, Stephen E.  
ALTUS, Christina M.  
DUFOUR, Gerard E.  
CHALUP, Michael S.  
HILLMAN, Jennifer L.  
JONES, Anissa Lee  
YU, Jimmy Y.  
WRIGHT, Rachel J.  
GIETZEN, Darryl  
LIU, Tommy F.  
YAP, Pierre E.  
DAHL, Christopher R.  
MOMIYAMA, Monika G.  
BRADLEY, Diana L.  
ROHATGI, Sameer D.  
HARRIS, Bernard  
ROSEBERRY, Ann M.  
GERSTIN, Jr., Edward H.  
PERALTA, Careyna H.  
DAVID, Marie H.  
PANZER, Scott R.  
FLORES, Vincent  
DAFFO, Abel  
MARWAHA, Rakesh  
CHEN, Alice J.  
CHANG, Simon C.  
AU, Alan P.  
INMAN, Rebekah R.

<120> MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS

<130> PT-1163 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/230,517; 60/230,599; 60/230,514; 60/231,167; 60/230,598;  
60/230,988; 60/230,518; 60/230,515; 60/229,751; 60/230,610;  
60/229,749; 60/229,750; 60/230,597; 60/230,505; 60/231,163;  
60/229,747; 60/229,748; 60/230,583; 60/230,519; 60/230,595;  
60/230,865; 60/230,951

<151> 2000-09-06; 2000-09-06; 2000-09-06; 2000-09-07; 2000-09-06;  
2000-09-06; 2000-09-06; 2000-09-06; 2000-09-05; 2000-09-06;  
2000-09-05; 2000-09-05; 2000-09-06; 2000-09-06; 2000-09-07;  
2000-09-05; 2000-09-05; 2000-09-05; 2000-09-06; 2000-09-06;  
2000-09-06; 2000-09-07

<160> 553

<170> PERL Program

<210> 1

<211> 1194

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:405741.3:2000SEP08

&lt;400&gt; 1

```

ccactgactg tgaaggggag gacgtggtgg acatgctcag ggaagccatc aagaggagaa 60
acgtaggatg tgggtgttgag gctcatgcct gctcttgctg ccttccccag gccctctgc 120
tgtgccatcc tctgcccagt tcctatgact ctggcccttc tccaggtcag gaggattctg 180
cagcagcttg gccctggaca gcacgtgtga ggatcactgt ggggtgtggac ggcaccctgt 240
acaagctgca ccctcacttt tctagaatat tgcaggaaac tgtgaaggaa ctagcccctc 300
gatgtgatgt gacattcatg ctgtcagaag atggcagtg gaaaaggggca gcaactgatca 360
ctgctgtggc caagagggtta cagcaggcac agaaggagaa ctaggaaacc ctgggatttg 420
acctgatgca tcttggatac tgaacagctt ttcctctggc agatcagttg gtcagagacc 480
aatggggcacc ctctgggtg acctcacctt ctggatggcc gaaagagaac cccaggttct 540
cgggtactct tagtatcttg tactggattt gcagtacat tacatgacat ctctatttgg 600
tatatttggg ccaaaatggg ccaacttatg aaatcaaagt gtctgtcctg agagatcccc 660
tttcaacaca ttgttcagggt gaggcttgag ctgtcaattc tctatggctt tcagtcttgt 720
ggctgcggga cttggaaata tatagaatct gcccatgtgg ctggcaggct gtttccccat 780
tgggatgctt aagccatctc ttatagggga ttggaccctg tacttgtgga tgaacattgg 840
agagcaagag gaactcacgt tatgaactag ggggatctca tctaacttgt ccttaacttg 900
ccatgttgac ttcaaactg ttaagagaac aaagactttg aagtatccag cccagggtg 960
cagagagggt gattgccagg gagcactgca ggaatcattg catgcttaaa gcgagttatg 1020
tcagcaccct gtaggatttt gttccttatt aagtgtgtgc catgtggtgg ggtgctgtct 1080
ggggcatctg tttttcattt tgctgtggtt ttgtgttgca ggtgttgata gttgttttaa 1140
ggattgttag gtataggaaa tccagtaaat taataaaaaa attttgattt tcca 1194

```

&lt;210&gt; 2

&lt;211&gt; 2324

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:337194.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 940, 971, 1034, 1088, 1107, 1158

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 2

```

caaacctcgt cccaactccc acccaccggc gtttctccag ctcgatctgg aggetgcttc 60
gccagtgtgg gacgcagctg acgcccgtt attagctctc gctgcgtcgc cccggctcag 120
aagctccgtg gcggcggcga ccgtgacgag aagcccacgg ccagagttaa tagagtggat 180
acaaccttgc tgaagatgaa gaatatcaa tattgaggat atttttttct ttttttttct 240
aagttcttga ttgtggctta cctcaagtta ccatttttca gtcaagtctg tttgtttgct 300
tcttcagaaa tgttttttac aatctcaaga aaaaatatgt cccagaaatt gagtttactg 360
ttgcttgtat ttggactcat ttggggattg atgttactgc actatacttt tcaacaacca 420
agacatcaaa gcagtgtcaa gttacgtgag caaatactag acttaagcaa aagatatgtt 480
aaagctctag cagaggaaaa taagaacaca gtggatgtcg agaacggtgc ttctatggca 540
ggatatgccc atctgaaaag aacaattgct gtccctctgg atgacatttt gcaacgattg 600
gtgaagctgg agaacaaagt tgactatatt gttgtgaatg gctcagcagc caacaccacc 660
aatggtacta gtgggaattt ggtgccagta accacaaata aaagaacgaa tgtctcgggc 720
agtatcagat agcagttgaa aatcaccttg tgctgtctca tccactgtgg attatatcct 780
atggcagaaa agctttataa ttgctggctt aggacagagc aatactttac aataaaagct 840
ctacacattt tcaaggagta tgctggattc atggaaactc aattctgtac ataaaaattt 900
taaagttatt tgtttgcttt caggcaagtc tgttcaatgn tgtactatgt ccttaaagag 960
aatttggtaa nttggttgat gtggaagca gatagggtgag ttttgtataa atcttttgtg 1020
tttgagatca agcngaaatg aaaacactga aaaacatgga ttcatctcta taacacattt 1080
atttaagnat ataacacgtt ttttggncaa gtgaagaatg tttaatcatt ctgtcatttg 1140
ttctcaatag atgtaacngt tagactacgg ctatttgaaa aaatgtgctt attgtactat 1200
attttgttat tccaattatg agcagagaaa ggaatatata tgttgaaaat aatgttttga 1260
aatcatgacc caaagaatgt attgatttgc actatccttc agaataactg aaggtttaatt 1320
attgtatatt tttaaaaatt acacttataa gagtataatc ttgaaatggg tagcagccac 1380
tgtccattac ctatcgtaaa cattggggca atttaataac agcattaaaa tagttgtata 1440
ctctaactct atacttattg aagaataaaa gatattttta tgatgagagt aacaataaag 1500
tattcatgat ttttcacata catgaatggt catttaaaag ttaatcctt tgagtgtcta 1560

```

```

tgctatcagg aaagcacatt atttccatat ttgggttaaat tttgctttta ttatattggt 1620
ctaggaggaa gggacttttg agaatggaac tcttgaggac tttagccagg tgtatataat 1680
aaaggacttt ttgtgctgca ttaaatgtct tggaaaatgt taacattata ttatataaga 1740
gtatccttta tgaatttttg aatttgtata acagatacat tagatattca ttttatataa 1800
tggccactta aaataagaac atttaaaata taaactatga agattgacta tcttttcagg 1860
aaaaaagctg tatatagcac agggaaacct aatcttgggt aattctagta taaaacaaat 1920
tatactttta tttaaatttc ccttgtagca aatctaattg ccacatgggt ccctatattt 1980
catagtattt atttctctata gtaactgctt aagtgcagct agcttctaga tttagactat 2040
atagaattta gatattgtat tgttcgatc tataaatatgc taccacatgt agcaataatt 2100
acaatatttt attaaataa atatgtgaaa tattgtttca tgaaagacag atttccaaat 2160
ctctcttctc ttctctgtac tgtctacctt tatgtgaaga aattaattat atgccattgc 2220
caggtagaag tttggaataa ttgttttagtc tctcctatta gatgtggaca tttttgtttt 2280
tgtttttggt ttcaggggat aaataaaata tattatttct acct 2324

```

&lt;210&gt; 3

&lt;211&gt; 615

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:017108.4:2000SEP08

&lt;400&gt; 3

```

cgagccgaga cggctacatg gacgccacta atcgccccgc accgtatccc ccccgagatg 60
ccccagtagc gggaggagaa ccacatcttc gaggtagatgc agaacatgct ggagcaactc 120
ctgatccacc agcccgaaga tcccatcccc ttcattgatcc agcacttgca tagagacaac 180
gacaatgtgc ccaggattgt aatattaggt ccacccgcct cagggaaaac aacaataggc 240
aatgtggctc tgcaaacatc tgaacagcag tctcctcacc ctggagaacc tgatcttaaa 300
tgagttttcc tatacgcca ccgaagccag aaggctttat ctgcaaagga agacagttcc 360
cagcgcgctg ctgctccagc tgattcagga acgctgggt gaagaggatt gcatcaagca 420
gggctggatt ctggatggca tccctgagac gcgtgagcag gctctgagga tccagacct 480
ggggatcaca cccagacacg tcattgtgct gagtgtccca gacacgggtc tgatcgagag 540
aaacttgggg aagagaatcg accctcaaac tggagagatt tatcacacca cctttgactg 600
gccacccgaa tctga 615

```

&lt;210&gt; 4

&lt;211&gt; 1837

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:372569.5:2000SEP08

&lt;400&gt; 4

```

cctccgcgca gtccttgagt cccgcaggcc ctgcgtcccc gctgcacacc cccgtccact 60
cccggtggtc ccggtccggc atggcgcgcg cggtggggcc cgagcggagg ctgctggccg 120
tctacaccgg cggaaccatt ggcatgcgga gtgagctcgg cgtgcttggt cccgggacgg 180
gcctggctgc catcctgagg aactgcccc tggtccatga cgaggagcac gcccgagccc 240
gcgccctctc tgaggacacc ctgggtgcta ccccgggcca gccgcaacca gaggatcctc 300
tacaccgtgc tggagtgcc gcccctcttc gactccagt acatgaccat cgctgagtgg 360
gttcgcgttg cccagaccat caagaggcac tacgagcagt accacggctt tgtggtcatc 420
cacggcacgg acaccatggc cttggctgcc tcgatgctgt ccttcattgt ggagaacctg 480
cagaagactg tcatcctcac tggggcccag gtgcccattc atgccctgtg gagcgacggc 540
cgtgagaacc tgctgggggc actgctcatg gctggccagt atgtgatccc agagggtctc 600
cttttcttcc agaactcagct gtttcggggc aaccggggca ccaaggtaga cgctcggagg 660
ttcgagctt tctgctcccc gaacctgctg cctctggcca cagtgggtgc tgacatcaca 720
atcaacaggg agctgggtgc gaaggtggac gggaaggctg ggctgggtgt gcacagcagc 780
atggagcagg acgtgggcct gctgcgcctc taccctggga tccctgccgc cctgggttcg 840
gccttcttgc agcctcccc gaaggcgctg gtcattggga ccttcgggtc agggaaacgga 900
cccaccaagc ccgacctgct gcaggagctg cgggtggcca ccgagcgcg cctgggtcatc 960
gtcaactgta cccactgct ccagggggct gtgaccacag actatgcagc tggcatggcc 1020

```

atggcgaggag	ccggcggtcat	ctcagggttc	gacatgacat	cggaggccgc	cctggccaag	1080
ctatcgatatg	tgctgggcca	gccagggtg	agcctggatg	tcaggaagga	gctgctgacc	1140
aaggaccttc	ggggggagat	gacgccaccc	tcggtggaag	agcgccggcc	ctcactgcag	1200
ggcaacacgc	tggggcggtg	ggtctcctgg	ctcctcagtc	tgagcggcag	ccaggaggca	1260
gatgccctgc	ggaatgccct	ggtgcccagc	ctggcctggt	gctgctgccc	acgccggtga	1320
cgtggaggcg	ctgcaggcgc	ttgtggagct	gggcagtgac	ctgggcctgg	tggactttaa	1380
cggccaaacc	ccactgcacg	cggccgcccc	gggggaggcc	cacacagagg	cagtcaccat	1440
gctgctgcag	agagggtgtg	acgtgaacac	ccggggacac	ggatggcttc	agcccgtgc	1500
tgctggccgt	gcggggcagg	catccgggtg	tcattgggtt	gctgcgggaa	gccggggcct	1560
cctgtccac	ccaggagctg	gaggaagcag	ggacggagct	gtgcaggcag	aaggcagccg	1620
ggaaacctgg	cagtgggtgg	ctttctacag	agcctggagg	gtgcggttgg	ttgccaggg	1680
ccccatgcc	agaagtgtct	ctgggtgtct	aacctgaagg	cgttctgctg	cagtataagc	1740
cattccttcc	tcccatgacc	tgctggaggg	gtctcaggca	tgacccact	gctggggctg	1800
cttcccagcc	tgctctcatg	taaagcctga	aggcctt			1837

&lt;210&gt; 5

&lt;211&gt; 709

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:968765.1:2000SEP08

&lt;400&gt; 5

gcctcagtgc	tttgcccgag	ctatgtgaag	gatcgctat	ggtgggagaa	gtggcggtga	60
cagttctgga	aagcagggca	gaggtgatag	ttcaagtaag	aaggacgtgg	tcgagctgac	120
ggatagacac	cttatgataa	gaatgtcctt	gacagtgaag	acgtttggat	ggttgagttt	180
tattgctcca	tggtgtggac	actgcaaaaa	cttggagcca	gaatgggctg	ctgcagccac	240
agaggtaaag	gagcaaacga	agggaaaggt	gaaattggca	gccgtagatg	ccaccgtgaa	300
tcagggttctg	gccagccggt	atgggattaa	aggattccct	acaatcaaga	tatttcagaa	360
aggcgagtct	cctgtggact	atgatggtgg	acggacgaga	tctgacatag	tgtcaagggc	420
cctgggatttg	ttctctgata	atgcgccacc	cccagagctg	ctcgagataa	tcaacgaaga	480
catagccaag	aagacatgcg	aggaacacca	gctctgtgtg	gtggccgtgc	tgctcaccat	540
cctggacacc	ggagctacag	gtagaaactc	ttacttggaa	gttcttctga	aactggctga	600
caagtacaag	aagaaaatgt	gggggtggct	gtggacggaa	gctggagccc	agtatgagct	660
ggagaacgca	ctagggatcg	gagggtttgg	gtaccctggc	atggcagca		709

&lt;210&gt; 6

&lt;211&gt; 533

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:255999.16:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 474

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 6

gatagggaca	agaagagata	gggctcgaag	ggtgtattgc	ctccaactgc	aaagcatttt	60
agacattggc	tggtctccag	gacaggaatg	tgcccaaaag	ggtgctttgt	atagtaccat	120
tcaggcttgt	ttctcaggtt	aatgttgacc	acttcatgtc	tgatgtcttt	ggccttgagg	180
acgaggcggg	tcctgtgaga	ataggggcag	aacctcatgc	tgtagatgcg	gatcagcccc	240
tccgggactg	gccctggggg	ctggcttcct	gggattttaga	gggttctgtg	ggcaaatgtg	300
ggagggaggg	cgcgggtggg	tctctggggg	gccaggcgtg	tgtaggactt	agctgtgagc	360
acacacggtc	ctgcccgaag	ggccaacgtc	tgggggatgc	agacaagagg	ctgggggtgga	420
gggtgctggg	ctcccggaca	gccccagggg	gtgtcggggc	tgtcaggaga	caanccacc	480
cccctctgct	gtgtggacca	ggacggggct	caggacagcc	gggctctctt	cca	533

<210> 7  
<211> 1428  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:977820.9:2000SEP08

<400> 7  
cgatcataat tcatagagta gaggcctggc tgggaggatc aatgatgcc aagtcaaggt 60  
ggttatcacc ttcaaccaag gactccgggg tgggcgcgtg gtggagctga agaaaatagt 120  
ggatgaggct gtgaagcact gccccaccgt gcagcatgtc ctgggtggctc acaggacaga 180  
caacaaggtc cacatggggg atctggacgt cccgctggag caggaaatgg ccaaggagga 240  
ccctgtttgc gccccagaga gcatgggcag tgaggacatg ctcttcatgc tgtacacctc 300  
agggagcacc ggaatgcccc agggcatcgt ccatacccag gcaggctacc tgctctatgc 360  
cgccctgact cacaagcttg tgtttgacca ccagccaggt gacatctttg gctgtgtggc 420  
cgacatcggt tggattacag gacacagcta cgtgggtgat gggcctctct gcaatgggtc 480  
caccagcgtc ctttttgaga gcacccagc ttatcccaat gctggtcggt actgggagac 540  
agtagagagg ttgaagatca atcagttcta tggcgcccca acggctgtcc ggctgttgc 600  
gaaatacggg gatgcctggc cgggcatggc caggaccatc tatggcgacc accagcgatt 660  
tgtggagccc tacttcaagg cctaccagg ctattacttc actggagacg gggcttaccg 720  
aactgagggc ggctattacc agatcacagg gcggatggat gatgtcatca acatcagtgg 780  
ccaccggctg gggaccgcag agattgagga cgccatcgtg agcacaggct gggcctctc 840  
cttcactcct acacctccat gatgcccctg tggacctgca agttggggtg ggagtcctc 900  
gcccagagtc cttctagcac ccggggccgc cacattttgc cttttgttc tctgtgttac 960  
acaggccgac caccctgcag taccagaaag tgctgtcatt ggctacccc acgacatcaa 1020  
aggagaagct gcctttgcct tcatttgtgt gaaagatagt gcgggtgact cagatgtggt 1080  
gggtgcaggag ctcaagtcca tgggtggccac caagatcgcc aaatatgctg tgcctgatga 1140  
gatcctgggt gtgaaacgtc ttccaaaac caggtctggg aaggctcatg ggcggctcct 1200  
gaggaagatc atcactagtg agggccagga gctgggagac actaccacct tggaggacct 1260  
cagcatcacc gcagagatcc tgagtgtcta ccagaagtgc aaggacaagc aggtctgtgc 1320  
taagttagct ggcaccttgc ggggtctctg ggatgggcgg gcacccaagc ctggcttgc 1380  
cttcccagaa ggtacccctg aggttggcgt cttcctacgt ccagaaag 1428

<210> 8  
<211> 543  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1071608.1:2000SEP08

<400> 8  
cggccgtcca gacgtcgtgg tggagtgaga gggtcccagg gaggaatgtg caaagtcctg 60  
ttggttctac ccctcttcta ctctcaggaa ggggtccgcg tcgcgcgttt ctggaacggg 120  
ctctctccct ggggtggctg ctgttctccg ccagattaag ccgggcctgc tggctcagta 180  
ctggagtcca gagctgaatt ctggccactt tgtcttctcc acagcctgtg ctcttgcca 240  
aggggacaat gacggcctcc agtgtcctcc tgcacactgg acagaagatg cctctgattg 300  
gtctggggac atggaagagt gaggctggtc aggtgaaagc agctattaaa tatgccctta 360  
gcgtaggcta ccgccacatt gactgtgctt ctgtatatgg caatgaaact gagattgggg 420  
agggcctgaa ggagagtgtg gggtcaggca aggcagtccc tcgagaggag ctgtttgtga 480  
catccaagct gtggaatact aagcaccacc ctgaggatgt agaacctgcc ctccggaaga 540  
cac 543

<210> 9  
<211> 673  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature

<223> Incyte ID No: LI:1074023.1:2000SEP08

<400> 9

```

gagatacttc atcgctccacg cagctttgag tccacacctc tgtctacgca gcagctatgc 60
cgccgtacac cattgtgtac ttcccagttc gagggcgctg tgaggccacg cgcagtctgc 120
tggtgaccca gggccaagagc tgggaaggagg aggtgggttac catagatgtc tggcttcaag 180
gctcgctcaa gtccacttgt ctgtatgggc agctcccca gtttgaagat ggagacctca 240
ccctttacca atctaattgcc atcttgaggc acctgggtcg ctctttaggg ctttatggga 300
aagaccagaa ggaggctgcc ttggtggata tggatgaatga tggggtggag gaccttcgat 360
gcaaatatgg tactctcacc tacaactaact atgagaatgg taaggatgac tatgtgaagg 420
ccctgcctgg gcatctgaaa ccttttgaga cctgctgtc ccagaaccag ggaggcaaa 480
ctttcattgt gggtaaccag atttcctttg cagattacaa cttgctggac ctgctgctgg 540
tccaccaagt cctggccccc ggctgcctgg acaacttccg ccctgctctc tgcctatgtg 600
gctcgctcaa gtgcccggcc caagataagg cttttctgtc ctctcctgac catttgaacc 660
gtcccatcaa cgg

```

<210> 10

<211> 691

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:453570.1:2000SEP08

<400> 10

```

cctgcacgct cctccgcctg attcccagcg cgtcgcactt caagagattc gatcgggttc 60
gccgggttcgc gcccgccgcc atggcgactt cctcaggacc aaaggaggcg ccagcgaata 120
accctgggct ccagaccgag gtcgaccccg ccaccaaagg ctacttcttg caacaaacga 180
tgttacgtgt gaaggaccca aaagtgagcc ttgacttcta ctacgtgtg atgggcatgt 240
cgttgttgaa aagggttgat tttgaagaga tgaattcag tttgtatttt ctagggttacg 300
aggatgtgac cttagccctt gatgatcata tcaagcggac tgaatggact tttaggcaaa 360
aagctactct tgagctcacc cacaactggg gcacggaaaa tgacctgaa ttcaaagggtt 420
accataatgg gaactcagac cctcgtgggt ttgggtcatat aggagtaact gtggatgatg 480
tccacaaggc atgtgaacgc tttgaacgct ttggggttga gtttgtgaag aaacctgatg 540
atggaaaaat caaaggcatc gctttcatca aagatcctga tggttactgg atcgaaatat 600
ttgaccagac gattgggacg gtaacttctt cagcatcatg atgaacagta gtggttcctt 660
atatggtgcg caagatcata caactctggc a

```

<210> 11

<211> 2930

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:072072.1:2000SEP08

<220>

<221> unsure

<222> 1826-1827, 1831, 1833, 1837, 1841, 1845

<223> a, t, c, g, or other

<400> 11

```

ggggggcgct gctacgggta tgagtccggc gtgcgcctcg tgacttttgc tcccacaagt 60
actgcctcgg aggcggggga gctggaccag cagaccgcct ggagcgcccg agtcaccgtc 120
gccggggctc ccgcgctccc cagaacgggtg gggactgcgg ggggttcggca gccgccagcg 180
gaacatggcg ccctggacgc tgtggcgctg ctgccagcgc tgtcgtgtgg ctgggtgccg 240
cgtgtctctt atcaccttcg tggctgtctg gtccctactac gcgtacgctg gtggagctct 300
gctgtgttta ctatttttgg aaagtgaagc aaaatggaaa gaccgttgtt taccttgttg 360
ctttccatct gttcttttgg atgtttgtat ggtcctattg gatgacaatt ttcacatctc 420
ccgcattccc cctccaaaga gttctacttg tccacttctg aaaaggaacg ttatgaaaaa 480
agaattcagt cctagaaaga caacaagaaa ttttcgagaa gagcagcaag agctttacct 540

```

```

atctataacca catcagcttc aaaaactatc agatattgtg aaaaatgtca gctgattaaa 600
cctgatcggg cgcacactg ctcagcctgt gactcatgta ttcttaagat ggatcatcac 660
tgtccttggg tgaataactg tgtggggatt ttctaattac aaattcttcc tgctgttttt 720
attgtattcc ctattatatt gccttttctg ggctgcaaca gtttttagagt actttataaa 780
atthttggacg aatgaactga cagatacacg tgcaaaatcc cactgacttt ttcttttctt 840
tgtgtctgca atgttcttca tcagcgtcct ctcacttttc agctaccact gctggctagt 900
tggaaaaaat agaacaacaa tagaatcatt ccgcgcaccc acgttttcat acggacctga 960
tggaatgggt ttctctcttg gatgcagtaa aaattggaga caagtctttg gtgatgaaaa 1020
gaaatattgg ctacttccaa tattttcaag cttgggtgat ggltgcagtt ttccaactcg 1080
ccttgtgggg atggatccag aacaagcttc tgttacaac cagaatgagt atgccagaag 1140
atctaaactg cagaggtaga aagcagaaaa cttggcgaat gacagtagtg gctcaaatca 1200
accttttctt atcaaaccac ttagtgaatc aaaaaaccgc ttgttggaca gtgaatctca 1260
gtggctggag aatggagctg aagaaggcat cgtcaaatca ggtgtatgaa aacattatag 1320
actggtattt tcaattttca tttgcaagaa aatgatcagt ggaatgaaat aactgaagta 1380
taacagaaga tatatttttt aaaacggaaa gcctttgtac agttcctggg attcacagaa 1440
gcactactcc agagcagaat gatgccttaa tcttaagtgt ccatttgtgc agcattgact 1500
tagagctacc aaagtgactt aatgttattc tggaaataat acttacctgt tatgagttgc 1560
tataatatga gctgtcatca catthttaaca tgcataatgta ttttttgata cctgaattac 1620
attattagaa taagtatcca tatacatttt cactccaaaa acacaagctt aaaaacttta 1680
aagtacatct agggcaaatg gtggctgaaa ttgaataata ttcaaattac tgttgcttgt 1740
actgtataag gaaaaactgt tttttaaatt gggatatttg taatttataa ttactatttt 1800
atacttttca acatttttaa ttagannctt ncnttgnttt ntttnccttt ttattgttgg 1860
ctataacaga aggattcaga tatctaaata tcatthtagaa catagaacac attggtctaa 1920
ttcatcgaaa gcttagttat gagcattaat acaaccatgt agtagagcaa gtgccaagg 1980
tttaaaggga acttggcagt ataagaaata ttttgggttc tctctgctat tttatagctt 2040
ttttttgttg aactttcagg ggtatgtaca tgggcagaga aacacaaagg tacatgtgct 2100
ggagaatact ggacacagta ggaagaagta taggtgtcca gaagaaaatc taagaggaat 2160
attgtacatt tgttaaagaa aaataagttt tgagthttta cagcacgtta gaataaagca 2220
tgctgcttga cttcacgttt gcaaatthca gtgtaggtaa tgactaaaaa ttgactttga 2280
aactttgtta cagtaatgac tgtaatagct taaacagtg cagtttaaat gcaattthaa 2340
ctaggcaaaa aaaaacttta gagtaagttt tccctcctgg actgtccagt tggcaaatcc 2400
tcaacctgca gtggctcatt ctacattccc agtctcttga gaagtggatt ctactcgta 2460
gacgcacagt cgtgaagtct gacttcctaa cggcaacccc ctttatgact agacacattt 2520
ctcctttgtc ccttattaaa atctaaggcg acaggaggtg ttgcattaa gctccaaggca 2580
aaagtacacg aataacacag atgctagaag agtgtcacgt cctcatttac acctthttaga 2640
cagacagaca cacacacaca cacacccccc ccaaaataat tgattgcttg ccagttagct 2700
gggtggagat gaactctcag tcagcactct cgtgcagtag ctacaggttg gagcctccct 2760
gtctgctcaa caatgggact gtcagggcgc agggcggtgc ttcagtgaac taagctcgga 2820
ctggatgact tctaacattc tgtgatgcct aattttgcaa aatcactttt cattcaccca 2880
ataaattttt ttcttctttt ttccacggag ttttgccttg tctcccagc 2930

```

&lt;210&gt; 12

&lt;211&gt; 1535

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:148565.4:2000SEP08

&lt;400&gt; 12

```

gcgctaaaaa cctctccacc gtccacctca gctcccatat ccgctccctt acctctccag 60
catctcccggt ctcgctcgte tcgtctccgg ctccctgcacc tcgcccgcag ccatgggcaa 120
gatcaagatc ggaatcaacg gtttcgggaag gatcggcagg actcgtggcc aggatcgccc 180
tgcagagcga ggatgtcgag ctcgctcgccg ctcaacgacc ccttcatcac caggattac 240
atgacactaga tgttcaagct acgacaccgt gcaacggcaa tggaaagcaca gcgacatcac 300
cctcaaggga ctccaagacg cttctcttct gcgagaagcc ggatcaccgt catttggcat 360
caggaacccc gagagatcc cgtgggggat aggtgtggct ctgagtatgt cgtggagtcc 420
accggtgtct tcaactgaaa ggacaaggct gctgacacat ctgaagggtg gtgccaagaa 480
ggttgttatc tctgacccca agacaaagga tgcgcccatt tatagtagtc aggtgtcaac 540
tgaggactgt tacactctcc gagggttaac attgttttcc aatgctagct gcacctacaa 600
actgctgtgt gctgccccgt gcctaaggct attcatgaca acttatggca ttattagagg 660
gcttgataga caacttgttt catgccatct actgccaccc tagaagtact gtttgtatgg 720

```



```

taccctcagt ccaaggactg gtagagtgtg gtcagggctg cccagcttta acatcattcc 780
cagacagcac cggatgctgg ccaaggctgt cgagtaagag ttcattccat gaattgaata 840
ggcaaagctc actgagcatg tccattccga gattcccacc gatggatgtg tcagttgttg 900
acctctactg tcagaatcga gaagtgggtg cctcttacga gtgaaatcaa cgaaagctat 960
taaggctgct tctgagtggc ccactcaagg tgtattatgg gcatacgtgg aaggaggatc 1020
tggtttccac cgaactttta ccggtgacat gcacgggtctg agcatctttt cgaaacgctc 1080
caaaggctcg cggatattgt cccctgtaaa cggagccaca cttttcatac aagacttcgt 1140
tcttctctgg tacgacaacg agtgggggta tcaacaaacc gcgttgctga gaccttgaga 1200
tctcgccgca atgtgtcaaa gaacccacgt atgagaagaa gaataatttc ttgccctccc 1260
ctatatcggt gggctctgtc cccacacaag tgcgcctttg tggttgacga accaactttg 1320
ctgcttgtct aatgctgaga ataaaaagtg agacgggtggc ccctcggaga cgctggagaa 1380
gcacagctgg ttttggagca acgcctgtct ttttgtgttt aaacttaatc tggctaaccg 1440
ctcctgtaaa cgaaattcac ctaagattct agttcttata tgctgctttg ttaagaaaaa 1500
taaacaagtg aggtttccaa aaaaaaaaaa aaagg 1535

```

<210> 13

<211> 1357

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:368626.4:2000SEP08

<400> 13

```

acacggatga ttcacacgta ggaatgcagg tgccctgtgc tcatctctgc ctgggaatgg 60
ggaggaacag atcttggggg acattgggga gtgggaggac aagcacccca ggagagatgg 120
tgaagtctct cactgcaatc atcttcaatt gctctgccca gcacgctgct gtcaacagtg 180
ggcagcatga ctttggggcc tggatgcccc atgctccatc atccattagg gaagcccca 240
ccccagacca aggggaccac caccctgaag acttacctag acaccctccc tgaagtgaac 300
atcagctgta acaacctcct cctcttctgg ttggttagcc aagaacccaa ggaccagagg 360
cccttgggca cctaccaga tgagcacttc acagaggagg ccccgaggcg gagcatcgcc 420
gccttccaga gccgcctggc ccagatctca agggacatcc aggagcggga acccagggtc 480
ttgcactgcc ttaaacttac ctggaacctt cccttcattg agaacagtgt tctccatcta 540
aaccaccccc aaataccacc caagaagaaa gaaaagggtc aagcatgaag aggaccagtt 600
ccctcaggtc ctccatagacc cttccatcct cctgttctt cagttcacct tgaaccttct 660
cttctgcaca ttggagactt ttgcagccaa gatggctctt gacatctata caaactgggc 720
cctgtagctg tgagtagtac cagcactagc agcgtccagg ttaaaagccg ctgtacctaa 780
agtccaatgc acttatagcc cctccgtaat aggaaggata ccgcttctac ttcttgcccc 840
tacttggggc aagcctcttg ttcttagcct ctgggcaagt gccagtcctt gggtttcatg 900
aagcatattt ctcccttcat cctcccccac ttccccaac tccttctcct accatgccat 960
ttatacgttc atctttcatt cacaagctct agtagtccat ccagcgccag cattacactt 1020
gactactggg aaaagcattg gaaaactagg gacaacagaa aggaactgtg tgccatgagg 1080
tctaacaatg tgggtcccat ttgtcacaca tagcaccatt taaggggagg ggaagaaatt 1140
ggagggcaga ttgcatggac cacaataacg gatgttataa ggaacaggat taaaaggaca 1200
atgggggtga ttgtggagat ctgtctttct ccaaagattt catccaactc ccctttcatc 1260
ccactccctt tcatcccat ttttatcttt ctgtccttga gccagtgag tttcaataaa 1320
aaccaaaata tttggctctc aaaaaaaaaa aaagg 1357

```

<210> 14

<211> 1116

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:346123.1:2000SEP08

<400> 14

```

ggtcgagcaa tctagaagca ccccatctcc catctctctc tcttcgccgt ccgaagctcc 60
gcaccccaaa tctaatacgaa caccctcacc agatggggcg caagtctctc gtcgggtggca 120
actgggaaat gcaatggaac cacagatcag gtcgaagaag attgtcaaaa ccctgaatga 180
aggacaggtt ccccttcag atgttgtgga ggtcgtttgt cagccctcct tatgtcttcc 240

```

```

ttcctgtggt caagagccag ctgcgccaa agttccatgt tgctgctcag aactgctggg 300
tgaagacagg gaggtgcttt cactgggtga agtcagtgc gagatgctcg tcaaccttgg 360
tggttccctg ggtcattctt ggacactctg taaaggagag ctctgctggg agaatacaat 420
gaatttggtt gagacaacgg ttgcgtaatg cctctgtctca gggcactaaa ggtcattgca 480
tgtgttggtg agacccttga gcagagggag gcttgggtct accatggatg tttgttgctc 540
gcacaaacaa agagcaattg ctgagaagat acacgggact ggagcaacgt agttgggttg 600
cctatggaac cagtttgggc tatattggaa ctggtaaagt tgccacctcc aggtccaggc 660
tcaggaagtg ccacgcctct cctgagggat tggctaaatg agccatatgc cagtccctga 720
ggttgctgaa ttctactagg atacatctac ggaggctctg taactgctgc gaactgcaaa 780
gagctagcaa gcacagcctg atgtcgatgg ttttacttgt cgggtggagct tctttgaagg 840
cctgagttca tcgacatcat caaacgcggc caaccgtgaa gtccgcttaa gatgctacgc 900
tgaaagacga acatactttt tttttgctca aactgtgcta tgataagcta gttagctatta 960
tgcgccaggag cagagactgt atttgccatg cccccaactt actagcttga gcttgctact 1020
aatgtttacc tctggacgta tcaataatgg tgcttatgta tccccctttt atcgtgtgccc 1080
aaaagagcaa cacaggggtcc ggcagcacia taattg 1116

```

&lt;210&gt; 15

&lt;211&gt; 3320

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:335795.11:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 1109

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 15

```

ccctcgccgc agtcgcgggc accccgctcg gcgtcggtgc ctgagggagg ccgcatgggc 60
ggccgaggcc ctaggcgggc gagggcgtgg cgtcgagcga ctggagcggc aggaggagga 120
cataccgctg gctgtggtcg gaggtcgagc gcctgagggg cgagcagctg aactcgccct 180
acagctgcca ggcggagggg ccgtgcctca cgcgggaggt ggccgagctc cgggccgaga 240
actgcgacct agcggccacc cctgtgcagc ctgcggctgt gcctcgccga ggagcggagc 300
cgccaggcca cgctggagag cgcggagcta gaggcggcgc aggaggccgg cgcacagcct 360
cctcctagtc aaagccaaga caaggacatg aaaaagaaga aaatgaagga aagcggaggct 420
gacagcgagg tgaagcatca accaattttc ataaaagaaa gattgaagct ttttgaata 480
ctgaagaaag accatcagct cttacttgcc atttatggac aaaagggggg atacaagcaa 540
catcatcaca gtaagagtgg ctgatgggca aacagtgcga ggggaagtct ggacaacaac 600
gccttacaca gtggctgctg aaattagtca ggaactggct gaaagcacgg taatagccag 660
agtcaatggt gaactgtggg acctggaccg cccattggaa ggggactctt ctctagagct 720
gcttacattt gataatgagg aagctcaagc tgtgtactgg cactccagtg ctcacattct 780
tggggaggcc atggagcttt actatggagg ccacctgtgc tacgggtccg ccattgaaaa 840
tggtttttat tatgacatgt tcattgaaga cagagcagtg tccagcacag aattgtcagc 900
cctggagaat atatgtaaag ccatcataaa agaaaagcaa ccttttgaaa gactagaagt 960
cagcaaggaa atcctcctgg aaatgtttta gtacaataaa tttaaatgcc gcattctgaa 1020
tgagaaagtt aacactgcaa ctaccaccgt gtacaggtgc ggtccattaa ttgacctttg 1080
caaaggtcca catgtaagac acactggana aattaaacc atcaaaattt ttaagaattc 1140
ctcaacatat tgggagggca atccggaaat ggaaacattg cagaggatct atggaatatc 1200
ctttcctgat aacaagatga tgagagactg ggaaaagtgc caagaggaag caaagaaccg 1260
agatcacagg aagatctggg aaggaacaag aacttttctt tttccacgat ttgagtcctg 1320
gaagctgttt tttccttccc agaggagcct tcatattata tacgcttaca gatttcatac 1380
gagaggaata tcacaaacgg gacttcacgg aggtgctctc tccaatatg tacaacagta 1440
aactctggga agcctcaggc cactggcagc attacagcga gaacatgttt acctttgaga 1500
ttgaaaagga cacttttgcc ctcaaaccga tgaatggtcc agggcactgt ctaatgcttg 1560
ccatcgtcc acgatcttgg agggaaactgc ctattagatt tgctgatttt ggagttctgc 1620
atagaaatga actgtcgggg actctcagcg gcttgaccag agtgaggcgc ttccagcagg 1680
acgatgctca catTTTTTTT cacagtggag cagattgaag aagaaataaa ggggtgtttg 1740
cagtttttgg catcctgggt tactcaacat ttggcttctc ctttcaatta aacctgtcaa 1800
caaggccgga aaacttccta ggcagagatt gacgatgtgg aactgaggct gaggaagcaa 1860
cgtgcagaac agcttgatgg acctttggag aaaccctgg caacaatggg gaaccaggg 1920

```

```

agaactggag gccatthttat tgggcccata aattgagcca taaaaatcaa gggatttgc 1980
aattggcaga ttaccatcc aatggtgcta caattccagc tgggacattc ccaactggct 2040
attagattta atctccacca ttattgttca gtcaaggatg gggatgataa gaagagacc 2100
tgtgatcatt ccacgagcc ccatttttgg gatccagtc gaacagaatg gcatagccat 2160
tgcttgctcg aaaacgtatg gcaggaaaat ggcctttctg gctatctcct cgtacagggtg 2220
atggtcatcc ctgtggggcc cacttgtgaa aacatattgca ccttcagggt tccagtgaat 2280
ttctttgagc gaagggattt atggctgacg ttgacttgga tcacacggtg gtaccctaa 2340
aataagaaca atacgcaatg ccagctggg ctgagtataa ttttatatta ggaggctgga 2400
gcacacagga cacgctagac tcatggctgt aaccgtgcga accaagagac aacaaaattc 2460
atgggagaga tttttagtaa cttctgccat tgataaactg aagaatctca cggaagacac 2520
ggacactcaa tgctgaggag gccttttgaa gtccttcctt gatattttgc ttctgtgtaa 2580
ctttgttttg acccctaaaa atgtattttt cttaacattg ttagtaactt ctacgacttt 2640
tggagccact gatggggtcc actcatggtc tcagctgaga aagtgagtag gatgtacgtg 2700
tagctgacat attgcacgaa gtttcaattt actcatgtcc acggggggacc gtttagaggg 2760
ccacgtggga aattttccag caatcaatgc cttgagaaac ttaaatgggg aatatattt 2820
catcgagaaa gtgaaacaaa aactaggaa atgacttatg aaactgttag gtgattttca 2880
aaagaatgtg gcttccaata aaagctctgt cagaggtttt ttttaacta ggagggaaa 2940
tcttattttc taggtatgtt ctgagggtat ttattatgac ttctaactaa aattcccacg 3000
tggaactttt tattcttcta aagctggaat catatactta aattttaact aaccaacaac 3060
cacaaaagcg agcagctact gcctaaatat tggattactg acaaaggaat tcagttttgt 3120
ggaatctggt gtttgacta taggttaaga gttgccattt aaatgtttct tattcataat 3180
taggttttgt tccgctttag aaaaaataa attcccaat gaattgcatc agtgtttcca 3240
gcgtttcagc tgtgggcttc taccaccact aagccgctca cctagagttg ttactttttt 3300
gatgggatga cattccatgg 3320

```

&lt;210&gt; 16

&lt;211&gt; 1281

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:246023.2:2000SEP08

&lt;400&gt; 16

```

ggcgggctcg ccaggacctg gcaaggcttg tttactatgg ccgatgatct ggagcagcag 60
tctcaaggct ggctgagtag ctggctgccc acgtggcgcc ccacttccat gtctcagctg 120
aagaatgtgg aagccaggat cctccagtg ctccagaata agttcctggc cagatatgta 180
tcctcccaa accagaataa gatctggacg gtgactgtga gccccgagca aaacgaccgc 240
accccttgg tgatgggtgca tggttttggg ggcggcggtg gtctctggat cctcaacatg 300
gactcactga gtgcccgccg cacactgcac accttcgac tgcttggctt cggggcgaagc 360
tcaaggccag cattcccaag ggaccggag ggggctgagg atgagtttgt gacatcgata 420
gagacatggc gggagacat ggggatcccc agcatgatcc tcctggggca cagtttggga 480
ggattcctgg ccacttctta ctcaatcaag taccctgata gagttaaaca cctcatcctg 540
gtggacccat ggggctttcc cctccgacca actaacccca gtgagatccg tgcaccccca 600
gcctgggtca aagccgtggc atctgtccta ggacgttcca atccattggc tgttcttcga 660
gtagctgggc cctggggggc tggctctggt cagcgattcc ggccggactt caaacgcaag 720
tttgacagct tctttgaaga tgataccata tcagagtata tttaccactg caacgcacag 780
aatcccagtg gtgagacagc attcaaagcc atgatggagt cctttggctg ggccggggcg 840
cctatgctgg agcgaattca cttgattcga aaagatgtgc ctatcactat gatctacggg 900
tccgacacct ggatagatac cagtacggga aaaaagggtg agatgcagcg gccggattcc 960
tatgtccgag acatggaggt taagggtgccc tcccaccatg tctatgctga ccagccacac 1020
atcttcaatg ctgtgggtgga ggagatctgc gactcagttg attgagctgc tctctgaaga 1080
ggaagaggag aaagccagag agtcaactct acctccctgt ctgcttactc acccactctg 1140
tcctttcctc accaactaac atgtgccagc caggcagagt cttgtgctgt tcccagaaca 1200
ggacgacagt gaaaagaaca ctcttgacc tacactgaag gctgaaggca gaagccacaa 1260
gaggccttga gtgccacccc c 1281

```

&lt;210&gt; 17

&lt;211&gt; 466

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1100661.1:2000SEP08

<400> 17  
atctcggact gaacaccaac gccatggcct gccccagctc cgcttgctgc ctgcttggcc 60  
tactggctct gacctccgcc tgctacatcc agaactgccc cctgggcccgc aagagggctg 120  
cgctagacct ggatatgcgc aagtgtcttc cctgcccacc cggcggcaaa gggcgctgct 180  
tcggggccgag catctgctgc gcggacgagc tgggctgctt cgtgggcacc gccgaggcgc 240  
tgcgctgccca ggaggagaac tacctgccct cgccctgccca gtctggccag aagccttgcc 300  
gaagcggagg ccgctgcgcc accgcgggca tctgctgtag ccggtatggc tgccgcaccg 360  
accccgccctg cgaccctgag tctgccttct ccgagcgtg agcccgcttt gtgatgatac 420  
ctttagggcg cttccttcat tccccatggc cactaccaga aaaaaa 466

<210> 18  
<211> 614  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> -Incyte ID No: LG:475856.1:2000SEP08

<400> 18  
attactcttg caggttccag atacatgaaa gaacccacac tggagagaaa ctttatgaat 60  
gtaaaacagt gtggtaaagc cttcagtagt tccagttata ttcatataca tgaacgaatt 120  
cactactggag agaaacccta tgaatgtaaa gaatgtggaa aacccttcag ttcccttaca 180  
ggcttttcgag tgcacatgag aatgcacact ggagagaaac cttataaatg taaggattgt 240  
ggcaacgcct tcatttggcg tgcacacctt caataccatg tgaagaaagt gcacgctgaa 300  
tagaaactat aaatgtaaat atcgaaagtt tttcaatttt aacgattttc aaagttgata 360  
tgaagattca tactggaggg tagttgcatg aatgtaagta atacaggaaa gcctgatata 420  
acttaattca ggtagaatgc tctagaaaat tcatgccatg aaaaaatatt ataatgtat 480  
ttatcagtg caccatccctt tttgattctt cctagttttt aatatatttt tatttttaaa 540  
tatttttata tatgtatata tatacgtata tattttgtaga gacaggcaac attatcttgc 600  
ccaggctggg cgcc 614

<210> 19  
<211> 615  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1015343.1:2000SEP08

<400> 19  
ggagacgcta ctggagtcgt ctctgctgcc tgcggacctg cgtctagact gacccaccgc 60  
gctcaagctc ggctgtctga ggcaggggag atgctgtcct gccgtctcca gtgcgcgctg 120  
gccgcgctct gcacgtcctt ggctttgggc ggtgtcaccg gggcgccctc ggaccccaga 180  
ctccgtcagt ttctgcagaa gtctctggcg gctgccaccg ggaaacagga actggccaag 240  
tacttcttgg cagaactgct gtccgagccc aaccagacag agaacgatgc cctggagcct 300  
gaggatttgc cccaggcagc tgagcaggac gagatgaggc tggagctgca gaggtctgcc 360  
aactcgaacc cagccatggc accccgggaa cgcaaagctg gctgcaagaa cttcttctgg 420  
aagacattca catcctgtta gctttaatat tggtgtctca gccagacctc tgatccctct 480  
cctccaaatc ccatactctt tccttaactc ccaggccccc cccaatgctc aactagacct 540  
tgcggttaag aaattgaaga ctgtcaatac caaattaaat tatggtgaaa ttatgcaaca 600  
accaaataaa agcgg 615

<210> 20  
<211> 964  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1400575.1:2000SEP08

<400> 20  
cgtcactgtg cgatcggggt gtgcttagct tgggggtctcc tggcccttg acgcgtcagg 60  
ttgctgtacc cctgcatcgg atgcgctgta ccctgcgctg gctccgtgaa ccttagggac 120  
aacaccggga caccgcgag gccggaaaat ggactcagt gcttttgagg atgtgtctgt 180  
gagcttcagc caggaggagt gggctctgct ggctccttca cagaagaaac tctacagaga 240  
tgtgatgcag gaaacattca agaacctggc atctataggg gaaaaatggg aagacccgaa 300  
tgttgaagat caacacaaaa accaaggacg aaatctaaga agccatacgg gagagagact 360  
ctgtgaaggt aaagaaggta gtcaatgtgc agaaaacttc agtcccaatc tcagtgtgac 420  
gaagaagact gccggagtaa aaccatatga gtgtactatc tgtggaaaag ccttcatgcg 480  
tctctcatcc cttactagac acatgaggtc tcacactgga tacgagctat ttgagaagcc 540  
atataaatgt aaggagtgtg agaaagcctt tagttatctc aaatcctttc aaagacatga 600  
aaggagtac actggagaaa aaccctataa atgtaaacaa tgtggaaaaa ccttcatata 660  
tcaccagccc tttcaaagac atgagcggac tcacattgga gaaaaaccct atgaatgtaa 720  
gcaatgtgga aaagctctta gttgttcag ttcgcttcga gttcatgaaa ggattcacac 780  
tggagaaaag ccctatgaat gtaaacaatg tgggaaagcc ttcagtgtgt ccagttctat 840  
tcgagtacac gaaagaactc aactggaga gaaaccctat gcatgtaagg aatgtgggaa 900  
agccttcatt tcccacacaa gtgttctaac acacatgata acacacaacg gagctgttcc 960  
agca 964

<210> 21  
<211> 672  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1080545.1:2000SEP08

<400> 21  
tattttctct tctttcaaat acatgaaagg aagcacagag aagagaaccg tatgcaatgt 60  
aagggttatg ggaaaacatt cagtttgccc agttttatttc atagacatga aaggactcac 120  
actggaggaa aaacctatga attgcaagca gtgtggcaga tcttcaact gttcgagctc 180  
ctttcgatat catggaagga ctcacactgg agagaaaccc tatgaatgca agcaatgtgg 240  
aaaagccttc agatctgcct cacagcttca aattcatgga aggactcaca ctggagagaa 300  
accttatgaa tgtaagcagt gtgggaaagc ctttggatct gcctcacacc ttcaaagtca 360  
tgggaaggact cacactggag agaaacccta tgaatgtaag cagtgtggga agtcttttgg 420  
atgtgcctcg cgacttcaaa tgcattggaag gactcacact ggagagaaac cgtataaatg 480  
taagcaatgt gggaaagctt ttggatgtcc ctcaaaccct cgaaggcatg gaaggactca 540  
cactggagag aaacctata aatgtaacca atgtggtaaa gtcttttagat gttcttcaca 600  
acttcaagtg catggaaggg ctcactgcat agacacccca taaccccagg ctttaggagg 660  
ctgagggtgg gg 672

<210> 22  
<211> 378  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:213947.1:2000SEP08

<220>  
<221> unsure  
<222> 10, 41, 44-45, 49, 61, 75  
<223> a, t, c, g, or other

<400> 22  
gccttctgan attgtgatga caacatagat agaaagcatc ntanntgcnt gcccatctgc 60  
nactgtttga ctatnaagct gtcagtggag tagaatatcg gaaatatttt catagaaatg 120

```

ttcgttggaa ttaatatattt ctttctggtg gtggcaacaa gaggacttgt cttaggaatg 180
ctgggaaacg ggctcattgg actggtaaac tgcattgagt gggccaagag ttggaaggtc 240
tcatcagctg atttcatcct caccagcttg gctatagtca gaatcattcg actgtattta 300
atactatttg attcatttat aatgggtattg tccccctcatc tatataccat ccgtaaaacta 360
gtaaaactgt ttactatt                                     378

```

&lt;210&gt; 23

&lt;211&gt; 1033

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:720641.1:2000SEP08

&lt;400&gt; 23

```

gacacttgcc caggtccctt gaataaaata caggtaaatac agtgcataag aagctttata 60
cagtgatctc agacttgtga ccaagagccc ctctgcacca ttgatagaca gtggccgtgc 120
ccaccaggag gatgggcaat cacactgcag tgagcatatt ccttctgtgg ggattttcca 180
gtttttcaga cctgcagagt ctactttttg tgggtattct cttctacatg tgaccatcct 240
agctgcaaac gtgtccataa tgggggccat caagctcagc cacaaccttc aactcctat 300
gtactttttc ctctgtggcc tgtccttttc agaaacttgt accactgtgg tagtaatccc 360
tcgcatgttg gtggactttc tatcagagag caagaccatt tctcttcctg agtgtgccac 420
acagatgttt ttctttctgg gctttgcac' caacaactgt ttcatcatgg ccgctatgtc 480
ctacgaccgc tacacggcca tccacaaccc actgcagtac cacacctta tgacaagaaa 540
gatctgcttg cagatgatga tggcttcttg gatggttggg ttctgtttt ctctgtgcat 600
catcgtcact gtattcaact tgtctctttg cgacttgaac actatccagc actatttctg 660
tgatatctca ccagtggtct ccttgcttg taattacact ttctatcatg aaatggctat 720
ttttgtgctc tctgcctttg tgttgggtggg cagctgtatt ttaattatga ttccctatgt 780
cttcattgtg ttcatagtca taaagatgcc ctctgcaaag gggagggtcta aggccttctc 840
aacttgctcc tcccacctca ctgttggtgc catacactat ggatttgctt gctttgtcta 900
tttgaggccc aagaacagca actccttcga tgaagacatg ctgacggcca tgatatatac 960
aatactgatg cctctgctta accccatcgt gtacagctct agaaacaaag aaatgcagat 1020
agccctaaga aaa                                     1033

```

&lt;210&gt; 24

&lt;211&gt; 963

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1023894.1:2000SEP08

&lt;400&gt; 24

```

ctggccactt cctgtcgccc ccaggcggac cctgcaggtc attcctgctc cttccctgcc 60
ctttcccact gtcacgtaac caacttcact agtcttcccc gaggggcaag agaaggatgc 120
ggcccacctg tgctatccag accccgatga cccaacagg gcggctgccg cccgcaacct 180
gcaactgctg gtgcacgacg tggattagga tgagttgccg ctaggggctc gggcccctgg 240
ggcaacgcgg gatgtagcga tacagaatgt tctgcgaga tgcaaggcca cttcctgtga 300
gcggctaccc atgttgaaac ggtatgaggg gttacttcct gtgtgacttg gggttgcttc 360
ctgtctctct gaaggccagg actttgccta gcaacatcac ttctgtgtg gggcgggact 420
aatttccatc atttcttttg cacagcccag accacgaccg tgagtgcctg aggtacgagg 480
ttgactgccc aagggtcgcg cgtgggctgt gcgctgggtg aggctgggac cctggcctcc 540
ctcatcacag gtgacgggtga acggcagtggt tggcgcccat ggggtccacg cgcccggtgc 600
ttgtaccgac gcggatatcg acatggctgc cgtggagggt ctggccccgc cgtgctgac 660
agctgaagtg cgaaagcacc ggggcccgtg tgaggtgggc gccgcagagc cggttccaga 720
cggccttcgt gttcactctg cagatcaacc agagttcaca gacagagcca aagcttgaga 780
tggtatatga gacagaactc aagggtcccga ccctggacac ggtttctgtg cgtgtcaagg 840
ccacagcctg gactcggggc tggagagcga ctggagcaag gcctggagcc tgggaagtgc 900
ctgctgagcc gcctgttccc acccatccaa aggatgagag tctacctg gccagagatg 960
gtt

```

<210> 25  
 <211> 1810  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:734904.1:2000SEP08

<400> 25  
 gacggcacct actgagaaat ttttactatc aacctaaaat ttaataagtc ctgggtcacaa 60  
 gtgaggaatt ttttaattgt tathtagatt atttttgtgt caaggtaact tgatcaccta 120  
 tcaataatgc ttataagaaa atagatatct catctgttgc tgtagaaaag gaatctgtct 180  
 cttccccaag tccttgccct caatgaggtt cccttgacca aaagtctggt cctcacattt 240  
 tagattcagc tttggaaatt ttgttcagga actgcatcat acaacccaaa agcaaaagag 300  
 ggaaaatgca taagccagga tacctaaatg ttcctttatt tagacttgaa aatgtgactg 360  
 caaaaggagt caagaagttt agtactcgaa cacattttctt cttttagtaa ataactgtct 420  
 aagaagataa attgttacct aacagttgat tgatgaatat gaaatcagag caatgagggg 480  
 ctaattagag attttactag ctagggtat ttaccataaa cctatctatc cagactgtag 540  
 acatagaatc accagatggc taggatccct ggtgcagggt gggtccctt tttccttacc 600  
 ctgaatgtca ttaaggatgc acttgccaaa tgctgccct ctggcctgat gactacactc 660  
 catattggct acctgccttc tttttcctac agtcttctc cagacaggca cgccatacaa 720  
 ccgactacac ttgggctcac tgaatgaatc acattcttct gctgtgcctc ccagagattt 780  
 catcaaaagca ccgcagtggt ccttctggag gctcctcaaa ctctccactc aatgtttcct 840  
 gaaagtgtcg tgtttctcac ctctagattg ttattctcat caagttacat tagtgaaatg 900  
 gtgaaaaatc agacaatggt cacagagttc ctctactgg gatttctcct gggcccaagg 960  
 attcagatgc tcctcttttg gctcttctcc ctgttctatg tcttcaccct gctggggaat 1020  
 gggaccatcc tggggctcat ctactggac tccagactcc acaccccat gtaattcttc 1080  
 ctctcacacc tggctgtcgt cgacatcgcc tacaccgca acacgggtgcc ccagatgctg 1140  
 gcgaacctcc tgcaccagc caagccatc tcctttgtg gctgcatgac gcagacctt 1200  
 ctctgtttga gttttggaca cagcgaatgt ctctggctg gtgctgaatg tcctacgatc 1260  
 gttacgtggc catctgccac cctctccgat actccggtca tcatgacctg gaaagtctgc 1320  
 atcactctgg ccataccttc ctggacatgt ggttcctcc tggtctctgag tccatgtggt 1380  
 tctcatccta agactgccct tctgtgggcc tcatgaaatc aaccacttct tctgtgaaat 1440  
 cctgtctgtc cacaggctgg cctgtgctga caccatggac tcaaccaggt ggtcatcttt 1500  
 tgcagcctgc agtgttctat cctgggtggg ccaccagac ctggtgcttg tcatcctact 1560  
 cgcacatcct ggcggccatc ctgaggatcc agttctgggg agaggccgca gaaaggactt 1620  
 ctccaacctg ctctcccaa ctctgacgtg agtgggactc ttctttggcc agtgccagtc 1680  
 atcatgtaca tggaacccca agtcccgaca tcctgaggag cagacaaaag gtcatttata 1740  
 ctatattaca gtttttttca aacccaacac ttaacctct gatttacagc ctgaggaacg 1800  
 ggcgtcag 1810

<210> 26  
 <211> 2620  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1178118.1:2000SEP08

<400> 26  
 ccacggggcg accgagatgg cggcctccgg gggcctagag ggaccctggg cgtgcactcg 60  
 cccttcgggc tcggccttta gttagtgaac agctcctcgg cgttctgcag agcgtgggtt 120  
 tcagcgagtt ctacgtgccg ggtccgccc gtgcggctt cctcgtgcc cctggcggt 180  
 cgtcagcccc cactaccct gaacttggtc ccaatggcgg ccgcccctc cttcaccgg 240  
 accgtgggca tctgggcctc gccgaagcgg tcaaggtggc tgctcgggct tctagagccc 300  
 gtgtccagcc ctttgccacc gaggcctgat cctcttttct gccctaaaga acttgccctg 360  
 acagcctctg gctcccgctc ttgaggatct tgcttgctca aaccacagaag acagtgcattg 420  
 aagccaggga catccgccat gctccaaaca acttggcctc agtagcggtt ccattcccca 480  
 aacctgctct gatctccac ctggagagag ggggaagcacc atggggccca gatccctggg 540  
 acaccgagat tctgagaggg atcagtcagg gtggtgagtc ctggatcaaa aatgaagggc 600  
 tagttataaa gcaggaagcc tctgaagaaa cagagttgca cagaatgcca gtaggaggac 660

```

ttctcaggaa cgtttctcag cacttttgatt ttaaaaggaa ggcactgaag cagactttca 720
atctaaatcc aaatctgata cttcgagggtg gaatgaagtt ctatgaatgt aaagaatgtg 780
ggaaaatctt ccgataatac tcaaagctta ttcgccatca gatgagtcac actggggaaa 840
agccctttta gtgtaaggag tgtggcacaag ctttcaagtc cagctatgat tgtattgtac 900
atgagaaaaa ccacattgga gaagggccct atgaatgtaa ggagtgtggc aaagggttga 960
gttccaacac agccttgact caacatcaga ggatocacac tggagagaaa ccctatgaat 1020
gtaaagagtg tggaaaggct ttccgtagga gtgcggcata cctgcagcat cagagattac 1080
acacgggaga gaaactctat aaatgtaagg aatgttgga agctttcggg tgtaggtcac 1140
tttttattgt ccatcagaga attcatactg gggagaaaacc ttatcaatgt aaggagtggtg 1200
gcaaaagcct caccagaaa atagcctcca ttccagcatca gagagttcac actggggaga 1260
agccttatga atgtaagggtg tgtgggaaag ctttcaaagt gtatggaagt tttgttcagc 1320
atcagaaatt gaccctgtg gagaagaagc cagtcaaggt ccttggggcca tccctgggtca 1380
gtccccagtg ctccctctcca gccatacctc ctgttcttct ccagggatcc tgttctgtct 1440
cagccgtagc tgtgccttca ctgaccttcc cacaatgctg gctcattcct acctctggga 1500
attttttcat gctgctgcct acatctggaa taccttcttc atctgccccaa atagtgcgtg 1560
tcttccaggg tcttactccc actgtgaaac cttccccagt tattctcacc ccttcttctc 1620
actcctcatg agctttatct tggcagtcct acggctctta tgcctagcaa atctccagcc 1680
taattttata tatttttttg agaaagggtc ttgctctgtc acccaggcta gagtgcgggtg 1740
gtgtgatctt ggtcactgc aacctcccc tctggggttc aagcgattct cctccttcag 1800
actctcgaat agctgggatt acaggcacgc ctccaccgc ccagctaatt tctgtatttt 1860
tagaagagat ggggtttcgc catcttggcc aggtcgtgtc caaacacctg acctcaagt 1920
atccgcctgc ctggccccc caaagtgtct ggattgcagg agtgagccac cactgtgcc 1980
agcctccaac ctaatttgag tttcatttgt atttttttca tcacctaca tgtgttaaca 2040
gctgtttcaa cataaactcc attataatct atatacattt aaagaccgtg tttacataat 2100
ctcttctggg tagaaaaatg aaatacttaa cctttgcatt ggtcccttcc aggcttgaac 2160
aacagtggct gtatctgcat cgtggggttg ctggtggtag aacactttgg tgttaggaaa 2220
gtttggaaaa gaggatatcc ctgtattagg ccattatgat tagagataag cagcctggga 2280
gaaacataca agacctgtgg ctactggtag agaattggaa taagcttcca acaatggact 2340
gtagagatta atggagattg tttagtgggc ccatggctag gtggggtatt aatacagttg 2400
aagttatgtt gcccttggca cagatctttg agaatctcat acatctgtga aaaagataat 2460
gtgaaataaa cagtaaaaaa aaaaagtatg ttgatggcag tgacgagtaa tcatgtcttt 2520
aaacgtgaag ttttcttgtt tctttagctt tggtctctga tagctagatg ttccctttaa 2580
cttcttattt cgtaagtaaa tatacacaaa ggcattcatc 2620

```

&lt;210&gt; 27

&lt;211&gt; 329

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:213947.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 12, 26, 46, 48, 53

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 27

```

tgcccatctg cnaactgttg actatnaagc tgtcagtgga gtagantntc ggnaatattt 60
tcatagaaat gttcggttga attaatattt tctttctggg ggtggcaaca agaggacttg 120
tcttaggaat gctgggaaac gggctcattg gactggtaaa ctgcattgag tgggccaaga 180
gttgaagggt ctcatcagct gatttcatcc tcaccagctt ggctatagtc agaattcattc 240
gactgtattt aatactattt gattcattta taatgggtatt gtccccctcat ctatatacca 300
tccgtaaact agtaaaactg tttactattt

```

&lt;210&gt; 28

&lt;211&gt; 845

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature



<223> Incyte ID No: LG:407304.1:2000SEP08

<220>

<221> unsure

<222> 823

<223> a, t, c, g, or other

<400> 28

atactttctc	agacaagcaa	aaaaatgaga	gttattgcc	gccaatctgc	cctacaaaac	60
atgttaaaag	cagctcttaa	gggagaaggg	aagagatata	ggtagaaaac	tcagatccac	120
ataaaggaac	tgtagagaaa	ggagtacatg	aagacaatag	aaaatctttt	acatttccta	180
ttcttagttg	atctagtagg	taactgcctg	gctaaagtaa	taatagtaac	attgtgttgg	240
ggtccatagt	aaaggatagg	tgaatcaag	aatagcactc	ttacaagggt	gaggagagag	300
gaactgggaa	tactctgcta	gaagacagct	gcactgcca	tgaagtggta	aagtgcatt	360
tgaaagtgg	tggtgattag	ttgcaaattc	atattgcgaa	ctctaggaca	atgaattaaa	420
atgttttgaa	agaagtataa	ttataatgct	agaaagtaaa	aagagtggaa	tcattgtaaa	480
ttctcagtta	aaccagaaaa	gacagaaaaa	gaaggggaag	taagaaaaac	aaggacaagt	540
gcacctgcta	gaaaacagcg	gcagacatgg	tagatattaa	tccaactatc	tcaataatca	600
ctttaaatgt	gaatggcgct	aaatatacca	actaaaacac	aaagactgcc	agtagattaa	660
aaaaaaatag	acccaatgat	atgttggtct	gtaaaaaaac	ccatttttaa	tataaagaca	720
caaataagcat	aaaaataaag	gtataaagaa	agatatacca	tgtaataact	aatcaaaaga	780
aaactggagt	agcagtattg	atatcagata	aggcaaaact	canagggaga	aaaatgatca	840
gggat						845

<210> 29

<211> 3427

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:337358.1:2000SEP08

<400> 29

gcgcggggag	aaagaagccg	tcctccggc	gcgtgggaga	gctgcagagc	cgcgggcaga	60
gaaggaggcc	gcggccgagc	ccggcagccc	gatacccccc	atccgagcga	ccccctttgg	120
ggcaggcgga	gaaggtggca	taggcgacca	ggagggttcg	cccacctggg	acacgcaccc	180
acagtcacct	ccaaaggacc	ccccctcccc	aagtctctag	gagctgctcc	ttcatcactg	240
gagtcgcccc	tacctctctg	ccccagccc	gagagcccca	ggcggggacc	cccggatcgg	300
acgtcccca	gcctccgggc	acctggctca	gcaggaggcc	cccggtcgg	ggcagggcag	360
ggcggcggg	ggcgagccgg	agcccgcccc	ctgcccgggc	cccgccgagc	cctcggagcc	420
cacccatggg	gcacctgccc	cttgccctc	cttgcccggc	cgcgcccagc	ccggcgctcc	480
gagcagcgca	ggggaggatc	cccgcgcagt	gacccgggag	ccaccacaga	ctctgggagg	540
ctcggcggt	ggagcagcag	gcagctcccc	gcagctcccg	gcgcttcag	gcagctctct	600
gagccgtgcc	agaggcccg	cccgccatc	ccagcccga	gcatgatga	agactttgtc	660
cagcgggaa	tgacagctca	gtgtgccgc	caaaaactca	taccgcatgg	tggtgctggg	720
tgctctcgg	gtgggcaaga	gtccatcgt	gtctcgcttc	ctcaatggcc	gctttgagga	780
ccagtacaca	cccaccatcg	aggacttcca	ccgtaaggta	tacaacatcc	gcggcgacat	840
gtaccagctc	gacatcctgg	atacctctgg	caaccacccc	ttccccgcca	tgcgagggt	900
gtccatctc	acaggggatg	tcttcacct	ggtgttcagc	ctggataacc	gggagtcctt	960
cgatgaggtc	aagcgccctc	agaagcagat	cctggaggtc	aagtctgcc	tgaagaacaa	1020
gaccaaggag	gcggcgagc	tgcccatggt	catctgtggc	aacaagaacg	accacggcga	1080
gctgtgccgc	caggtgccca	ccaccgagcc	cgagctgctg	gtgtcgggcg	acgagaactg	1140
cgctacttc	gaggtgtcgg	ccaagaagaa	caccaacgtg	gacgagatgt	tctacgtgct	1200
cttcagcatg	gccaagctgc	cacacgagat	gagccccgcc	ctgcatcgca	agatctccgt	1260
gcagtagggt	gacgccttcc	accccaggcc	cttctgcatg	cgcccgctca	aggagatgga	1320
cgctatggc	atgggtctcg	ccttcgcccc	ccgccccagc	gtcaacagtg	acctcaagta	1380
catcaaggcc	aaggtccttc	gggaaggcca	ggcccgtag	agggacaagt	gcaccatcca	1440
gtgagcgagg	gatgctgggg	cggggcttgg	ccagtgcctt	cagggagggtg	gccccagatg	1500
cccactgtgc	gcattctccc	accgagggcc	cggcagcagt	cttggttaca	gaccttaggc	1560
accagactgg	aggcccccg	gcgctggcct	cgcacattc	gtctgccttc	tcacagcttt	1620
cctgagtcct	cttgctccaca	gtccttggtg	ggtttcatct	cctctgtggg	aggacacatc	1680
tctgcagcct	caagagttag	gcagagactc	aagttacacc	ttcctctcct	gggggttgaa	1740

```
gaaatgttga tgcagaggg gtgaggattg ctgcgtcata tggagcctcc tgggacaagc 1800
ctcaggatga aaaggacaca gaaggccaga tgagaaaggt ctctctcttc ctggcataac 1860
accagcttg gtttgggttg cagctgggag aacttctctc ccagccctgc aactcttacg 1920
ctctggttca gctgcctctg caccctctcc caccctcagc acacacacaa gttggccccc 1980
agctgogcct gacattgagc cagtggactc tgtgtctgaa gggggcggtg ccacacctcc 2040
tagaacacgc ccaccactta gaccacgccc acctctgac gcggttcctc agcctcctct 2100
cctaggtccc tccgcccagc agttgtgctt gtttgggtt gcagctgttt tcgtgtcatg 2160
tatagtagta gaaatggaaa tcattgtact gtaaaagcct agtgactccc tccttggcca 2220
ggccctcacc cagttcagat ccacggcctc caccgggac gccttctctc tctgtctcca 2280
aacagggttt ccgtggcctg tttgcagcta gacattgacc tccgccattg agctccacgg 2340
tttacagaca attgcacaag cgtgggggtg gcaggccagg actgcttttt tttaatgctc 2400
ccatttcaca gaggatacca ccgagactcg gaggggacac gatgagcacc agggccccacc 2460
tttgtccctt agcaaattca ggttacagct ccacctagaa ccaggctgcc ctctactgtg 2520
ctcgttcttc aagcatttat taagcaccta ctgggtgctg ggttcactgt gtcttaggaa 2580
accaagaggg tcccccagtc tggcctctgc ccgcccctgc tgccccacca ccttctgcac 2640
acacagcggt ggggagggcg ggaggagcag ctgggaccca gaactgagcc tgggagggat 2700
ccgacagaaa agctcagggc ggtctctctc ccttgtgccc gggattgggc tatgtgggt 2760
accacatgt actcaggcat ggtgggtttt gaaccataa accaaaggcc cttgtcatca 2820
gctcttaaca agtatatttt gtattttaat ctctctaaac atattgaagt tttaggggcc 2880
taaggaacct tagtgatctt ctattgggtc tttctgaggt tcagagaggg taagtaactt 2940
cctccaggtt cacacagcaa gtctgtgggt ggcagaagca agctagcact gggcattcag 3000
tacataccac gatgtgctcc ctctcttgat gcttggcccc tggggccttc agggcttttg 3060
gacatcttgt cctcaacctt ctccctagat cagtctgtga gggctccctg agatattgtg 3120
tacacatgc ccatgtatat acaagtacac acagatgtac acacagatgt acacatgctc 3180
cagccccagc tctgcatacc tgcacctgca cccagcctt ggccccctgc tgcgtctgtg 3240
ctcaaagcag cagctccaac cctgcctctg tccccctccc caccactgc ctgagccttc 3300
tgagcagacc aggtaccttg gctgcaaccg tgtgtggccc gctctcacc aggcacagcc 3360
ccgccaccat ggatctccgt gtacactatc aataaaagtg ggtttgttac aaaaaaaaaa 3420
aaaaggg
```

&lt;210&gt; 30

&lt;211&gt; 831

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:986090.1:2000SEP08

&lt;400&gt; 30

```
ccgcagttcc ccgaccacac cgttctctcc agcccaggag cagccccgcg ccaggaggaa 60
gtttgttctt ttaacaatat gtcgagggag gagaatgttt acatggccaa gttggctgag 120
caggctgaaa ggtatgagga gatgggttag tatatggaga aggtggctaa gactgtagat 180
gtggaagagc tcaactgttg ggagcgtaac ctctctgtctg ttgcatacaa gaatgtgatt 240
ggggctcgcc gcgcttcatg gcgcattggt tcttccattg aacaaaagga ggagtcccg 300
aagaacgaag agcatgtgaa cttatcaag gaatatcgtg gcaagatcga ggctgaactg 360
agcaatatct gtgatggaat cttgaagctg cttgactccc acctagtgcc ttcttctact 420
gctgctgaat caaaggctct ctacctcaag atgaagggtg actatcacag gtatcttgcg 480
gaattttaaga ctggtgctga gaggaaggaa tctgctgaga gcacaatggt ggccataaag 540
gctgctcagg acattgctct ggctgagctg gcacctaccc atccgataag gctgggactt 600
gctcttaatt tttcagtggt ctattatgag attctgaact cccagacaa agcttgcaac 660
ctggcaaaagc aggcgtttga tgaagctatc tctgagttag acagccttgg tgaggagtca 720
tacaaggata gcactctgat catgcaactc ctgagggata acttgaccct tcggacatcc 780
gacctcacgg aggatgggtg tgatgagggc aaagatgccc cgaaagggtga t 831
```

&lt;210&gt; 31

&lt;211&gt; 1364

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:123250.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 755

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 31

caagcaccga	ggaggcccca	gctccctagg	ggctgagaag	ctggagtctg	ggcaagggga	60
ggagctgagc	cctactcttg	caagaccccc	ggcctcctca	ccccacgcgg	gaagcatgaa	120
cagaaaagac	agtaagagga	agtccaccca	ggaatgcacc	ggaaaaacag	gagggcgagg	180
ccggcccccgc	caagtgcgcc	gccacaagac	atgccccagc	cctcgggaaa	tcagcaaggt	240
catggcttcc	atgaacctgg	gcctgctgag	tgagggcggc	tcagcgaag	atgagctgct	300
ggagaaatgc	atccagtcct	tcgattcagc	tggcagcctg	tgccacgagg	accacatgct	360
caacatggtg	ctggccatgc	acagctgggt	gctgccgtcc	gccgacctgg	ctgcccgct	420
gctgacctca	taccagaagg	ccacagggga	cacccaggag	ctgagacggc	tcgagatctg	480
tcacctgggtc	aggtactggc	tgatgcgaca	ccctgaggtg	atgcaccagg	atccccagct	540
agaagaagtc	ataggtcggt	tctgggccac	cgtggccccg	gagggcaact	cagcccagag	600
aagactggga	gactcttctg	acctgtgagt	cagccctgtg	ctccccctct	ccctgccccca	660
tcgccattct	gccatctccc	tctatatccc	caagccctca	actgtgagcc	actccaatac	720
cctgccccctc	ttggccccca	gcctgagccc	tggtngccct	ggccccccac	tcccaatgag	780
cagcccaggc	ctgggcaaaa	agcgcaaaagt	gtccttgctt	ttcgaccact	tgagacggg	840
ggagctgggt	cagcacctca	cctacctgga	gttccgggtc	ttccaggcta	tcacgccccca	900
ggacctgcgg	agctacgttt	tgacgggctc	agtacgaggc	tgccccgcct	ggagggctcc	960
gtaggtctca	gcaacagcgt	gtcccgcgtg	gtgcaggtga	tggtgctgta	gccgtccccg	1020
gccctacag	cgtgcacagg	tgctggacaa	gttcattcac	gtggcacaga	ggctccacca	1080
gctgcagaat	ttcaacacgc	tgatggcagt	cacagggggc	ctgtgtcaca	gtgccatctc	1140
cagactcaag	gactcccatg	cccacctgag	ccctgacagc	accaaggccc	tcttgagct	1200
cactgagctc	cttgccctccc	acaacaacta	cgcccgctac	cgccgcacct	gggctggctg	1260
cgcgggcttc	cggctgtttg	tactgggcgt	gctactttaa	ggatctgggt	tccctgtatg	1320
aggtatagtc	ttgacaggtt	gcctgacggc	cgctgcacc	aacc		1364

&lt;210&gt; 32

&lt;211&gt; 1614

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1028774.2:2000SEP08

&lt;400&gt; 32

tgcattgatg	gtttagatgc	atacaacttt	gaagtcaggt	tgtataacctg	tgaagaagtg	60
gacagagaca	aggctacccc	taccatattc	gccacagccc	caagaataga	cttccttggt	120
tcgtgtcaga	accaccacat	gattatctcc	acaggagacc	tgctccactg	gattgtctgag	180
gaagaagtcc	agctgcatgg	gttctagcac	ttcagggcca	gcaactttgt	ccacccccat	240
gcagccataa	taatctgata	cgaaggcata	gagctgaccc	tcatcagtc	cacagacagt	300
gaaatcatca	ccacatgaca	cctgatggat	agctttgcct	tgcaactttt	ccacatgctt	360
tggtgtctga	taggaggctt	tgtctccatg	gccagctga	ccatggagtt	tagtgcctcc	420
ttgcatgttc	acccaagtgt	acagttcctt	ctccactgtg	accacagcaa	agtgggtatt	480
ccctgcacag	acctgccggg	cactacagcc	actcttgata	acatccagtt	tctggggggg	540
ggatttttcca	ccacccccaa	catagacttc	actgggttcgt	gatgttacta	cagcaatggg	600
tgtctcagtc	acagtgcctg	accttgggtc	ctttgtaggt	gcattaagca	gagtgcactt	660
ttctctccatc	tctctcctgc	gtttcctgag	aagagggcga	tctagaagtt	catctgcagt	720
aggtctctgc	tcaggatcct	ggtcaaggca	cgaatgaacc	atttggtatca	attccaaaga	780
gtactgggcta	gagtcaactt	ccatggcccc	aattccttgc	acgatcttca	cacacaggtt	840
aagtgggttt	gtagcatcaa	acgtcctctt	taaggtaagc	agttcaaaaa	tgacgcagcc	900
aactgcccag	atatcagact	tgaaattgta	ctttactcct	tgacagagct	ctggagacat	960
gtaatatggg	gttcccacaa	gcgtctcagc	catggaatac	tcagaattaa	gtttcttttgc	1020
tagggccataa	tctccaagtt	ttatcaggtt	tgcttgggtc	agaaaaatat	ttaatgtctt	1080
tatatctcta	tgaaggattc	cagctttatg	gatgcagctc	actgctgaaa	caatctgaaa	1140
taggtaccac	accaccatct	cttctcaaaa	caacttgtcc	ttctgacgaa	ggattttgtc	1200
atacagggtc	cctccattac	aattattccag	ctcaatcagc	agcgtgggtat	tgctccatgaa	1260
gtgattgtag	taggcaataa	tggtgtcatg	ctgcagcagt	gccagaataa	ctatctcatt	1320
caaggcatca	cgacgttcct	tctcagacag	ccgggtcaaa	tcgacttcct	tccacacaac	1380

```

cagtgaagtca tcctcgggtgc ggcgggtacag cgtgggcttcc ccgaaggcgc cgccggccag 1440
gacgcgggatg gggatgtagt gcagttcctc ctgctccgcc gcgccgccgc cggctcgcgg 1500
ccctctgactg gcgctaggcc ccggactcga gtccccgcaa ccccgaggact cgctcccaaa 1560
gtccgagttg atggaatcgc agtgtcgctc gtactcgccc agcaccgaca tggc 1614

```

&lt;210&gt; 33

&lt;211&gt; 1254

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:338927.6:2000SEP08

&lt;400&gt; 33

```

cccacctaca ggttgtgaga attacatgga atgagacatg catcatgtcc tgggtagtgc 60
ttggtaggta gccctcgta ttgttacata atgaagcact gagcagttac tgaacacctg 120
ctgtggggccc aacactgtgc cgggccttgt ggatgaggcc aacctgggac cagaacagcc 180
tttggcctca ggaggctcac agtctaagga gtagaggggc gcatgtgagg atgacacccc 240
cgtacccccc ccacacccag aattgccgtg gaaagtttgc tttgggggat tcaggagacc 300
tggcctcacc tggccctatc tagagtcgag aggaggtctc taggctgagg gcaatccctg 360
tgccctcctc gcaggacagc taacccccgc atcggtgacg tgatccagaa gctgggcccc 420
ttcctgaaga tgtacagtga gtatgtcaag aactttgagc gagcggctga gctgctggcc 480
acctggaccg acaagtctcc actcttcag gaggttctca ctgcgcatcca ggtgaggctg 540
ggggaggggct ggagtcagca ttgccactcc cagcatgcag tggctcaggt tgccttgagt 600
gattccgggc atctcccagg ctccagctgct tccataggcc cctgcctact cgtccggccc 660
tcaggagcag cctgacccac ctcccttctc tcacctctc cgtgttgcct ccccatccct 720
cccaagagc agcgaggcct cgggcagcct gacctgcag caccacatgc tggaaaccag 780
gcagagaatt ccaggttacg agctgctgct caaggagtag atccagaagc tggcagccca 840
ggccccagac caggccgatg ccagaaaagc cctggacatg atcttctcag ctgcccagca 900
ctccaatgca gccatcactg agatggagcg gctgcaggac ctgtgggagg tgtaccagcg 960
cctgggcctc gaggacgaca tagtagaccc ctctaaccac ctgctccgtg agggcccggg 1020
cctcaagatc tccttcggcc gcaacgaccc catggagcgc taccttttct tgttcaacaa 1080
catgctgctc tactgtgtgc ccagggtgat ccagggtggc gccagttcc aggtgaggac 1140
ccgcatcgat gtggccggga tgaagggtgc ggagctgat gatgctgagt tccccactc 1200
cttctcgggtg tctgggaagc agcgcaccct ggagctgcaa gcccggtccc agga 1254

```

&lt;210&gt; 34

&lt;211&gt; 2736

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:332944.2:2000SEP08

&lt;400&gt; 34

```

ctgcgagctg ccggggcactc agcgcgggtc atggcgtgga tactggactg ccttttccgc 60
tcggcctttg agccccgccc ccgcgctgtg agtgtcttgg gaggagcccc aggacacaac 120
ccgaccgca ggacgaagat ggtatcgata cacagcctct ctgagctgga gcgtctgaag 180
ctgcaagaga ctgcttaccg cgaactcgtg gccagacatt tcctctccga attcaaacct 240
gacagagctc tgccatttga ccgtccgaac accttgata agtggtttct gattttgaga 300
ggacagcaga gggctgtatc acacaagaca tttggcatta gcctggaaga ggtcctgggt 360
aacgagttta cccgccgcaa gcatcttgaa ctgacagcca cgatgcagggt tgaagaagcc 420
accggtcagg ctgcggggccg tcgtcgggga aacgtgggtg gaagggtgtt tggccgcata 480
cggcgctttt tcagtcgcag gcggaatgag ccacacttgc cccgggagtt cactcgcgt 540
gggcgtcgag gtgcagtgct tgtggatagt ctggctgagc tgggaagacgg agccctgctg 600
ctgcagaccc tgcagctttc aaaaatttcc tttccaattg gccaacgact tctgggatcc 660
aaaaggaaga tgagtctcaa tccgattgag aaacaaatcc ccagggttgt tgaggcttgc 720
tgccaattca ttgaaaaaca tggcttaagc gcagtgggga tttttaccct tgaatactcc 780
gtgcagcgag tgcgtcagct ccgtgaagaa tttgatcaag gtctggatgt agtgctggat 840
gacaatcaga atgtgcatga tgtggctgca ctcccaagg agtttttccg tgacatgaag 900
gattttctgc tgccagatga tctgtacatg tcattcctcc tgacagcaac tttaaagccc 960

```

```

caggatcagc tttctgccct gcagttgctg gtctacctga tgccaccctg ccacagtgat 1020
accctggagc gtctgctgaa ggccctgcac aaaatcactg agaactgcga ggactcaatt 1080
ggcattgatg gacagttggg ccagggcaac cgtatgactt ccactaactt ggccttggtg 1140
tttggatctg ctctcctgaa aaaaggaaag tttggcaaga gagagtccag gaaaacaaag 1200
ctggggattg atcactatgt tgcttctgtc aatgttggtc gtgccatgat tgataactgg 1260
gatgtcctct tccaggtgcc tccccatatt cagaggcagg ttgctaagcg cgtgtggaag 1320
tccagcccgg aagcaactga ttttatcaga cgcaggaact tgaggaagat ccagagtgca 1380
cgcataaaga tggaaagagga tgcactactt tctgatccag tggaaacctc tgctgaagcc 1440
cgggctgctg tcttggctca aagcaagcct tctgatgaag gttcctctga ggagccagct 1500
gtgccttccg gcactgcccg ttcccatgac gatgaggaag gagcgggtaa ccctcccatt 1560
ccggagcaag accgcccatt gctccgtgtg ccccgggaga aggaggccaa aactggcgctc 1620
agctacttct ttccttagat gtttttcctt ctataaggtg ccagacaggg gaaaagggtg 1680
ggggtacatc tgggatgtca caggaaacat taaggagaga gttgaaggta aagatctgaa 1740
ccctaagaag agttccacct gatgctcggg tcaggatgag aattccaaac acactgccag 1800
cccttccact ggggatgctt ggtctcttct gctggtaaaa gcagagatgt ttctgtgtca 1860
tgcccaagct ccccggtgct accttgccct tctcttttac cctgatctt ggctttctct 1920
ctctctctgc agactttcct ttaattgatg tgacatttgt ggtaaacacc ttcccagggt 1980
aacctcacia atcttgagat gctttccctt ccccaaatgg gattgcatga tttccctgac 2040
tttccctacc tccctcagag agctcagttg gaaaggccct caagaggcat gctagaacgt 2100
taggtcagcc tactgacagc tgacaaacaa ttaatgcgaa atcatgtcac accaaccat 2160
agccgtgtcc acgcagcaac tccaccacct taggatttcc cctccaaat tattcagacc 2220
aatggcttgc caaatggcct ctcccaaaat tctgtacagt tttgctcagg tcacgccaac 2280
agggaaacct caagtgtagg tctaattagt gtttctggga tccaaagtta gaggaaaaat 2340
tagattttat tgcctggatc tgctttaaag acaattgggt ttacaccct ctgtcagca 2400
aaacagctag ttaggtaagg acatatagtt ccaagtaggt aaagtcaact gattacaaat 2460
gttcttaact atcgtctctg taattccttt atacaggaca gtacaaaatt gtgggacatg 2520
ctctggtaac acacagatat gggttgcata tgatccagaa ttacagctga tattatggat 2580
gacaactgct aagggtccata aaatgaagac tgtattgtat tgagggatag aaattgatca 2640
tttaatgggt aacaactgct gagctcaaag atttgtgatt gttaaaactt ctctggcatt 2700
taatcattaa taaacatctg tattgtgaca gcagca 2736

```

<210> 35

<211> 1518

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:347174.5:2000SEP08

<400> 35

```

aatgcagccc cccaggccgg ccgacttcaa gttgcaggtc atcattatcg gctcccgcgg 60
cgtgggcaag accagcctga tggagcgctt caccgacgac accttctgcg aggcctgcaa 120
gtccaccgtg ggtgttgact tcaaaatcaa aactgtagag ctaagaggaa agaaaattag 180
attacagatc tgggacacag caggtcagga gagattcaac agcattacct cagcttatta 240
cagaagtgcc aaggggatca tattagtata tgatatcact aagaaggaga catttgatga 300
tttgccgaaa tggatgaaga tgattgataa gtatgcttca gaagatgcag agcttctctt 360
agttggaaat aagttggact gtgaaacgga cagagaaatc accaggcagc agggggaaaa 420
gtttgcacag cagatcactg ggatgcggtt ctgtgaagca agtgccaagg ataacttcaa 480
tgtggacgag atatttttga aacttgtcga tgacattctg aaaaagatgc ctctggatat 540
tttaaggaat gagttgtcca atagtatcct gtcgttacia ccagagcctg agataaccgc 600
agaactgcct ccaccaagac cacatgtccg atgctgttga tttcctactt tggagacaaa 660
gtggaaatga ttcctggaaa ggggaaaaaa cgttctattc tgcactacaa tcattttgac 720
aatttccttt cgcactttgt aatccaagtc agagctatac actaacttgt aaatatgcat 780
atatgcaatc ctgggtaagt tttggttata agttacctat ttccctccaa attattatat 840
ttcattcatt actccagtgt ctagtgtaca tacactggga aacctagtac ttctaatatg 900
aagaatggga gaaatgaaag gtataatggt tcttgaaata aataatataa ttgtccttat 960
taattatatt atgaggacag aagatattct gataagagag aacgtggtgc tttgcttacc 1020
gttttaaaaga aaatttgtaa aactaaagac tttttgaaaa aaagctatct taagtgtctt 1080
ttcttttatt acaagacatt tccccagtg tagcatctg aagtattgga gtgtttctgc 1140
cacgaagcaa agctccattc atggccgtca tgggaaggta tttaataatg ttacataatg 1200
gtagaatatt actagttaga gggttggatt tgacttggtc ctaaggccac agaattctctc 1260
tcatggcttc ctaagggatg tacctttatg cttttaagaa ctacaaagat tcaataaaga 1320

```

```

aagaaatggt tttgaaacta tagaaaaaga ttttaaaaca cgctgctgtc ctaaacaat 1380
cctgttttaa ggaattttta agagatgcat tttactatat caaagaacat acgtgtattt 1440
gcctaaacac tctgtacctt tgtaatgata aaacttcccc cttctttacg gtgaagctta 1500
ttctgattaa gcctagac                                     1518

```

&lt;210&gt; 36

&lt;211&gt; 798

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:477070.1:2000SEP08

&lt;400&gt; 36

```

gaaaaacttt tctgcaatga acaagctcaa gaagatggca ttgaggggtga ttgctgaaag 60
tctatctgag gaggagattg ggggtttgaa ggagttgttc aaaatgatcg atactgacaa 120
cagcgggacg ataacttatg atgaactgaa ggatggctctg aaaaggggtg gctcggacct 180
gatggaacct gaaatccagg ctttaatgga tgcagctgat atcgacaaca gcggaacct 240
tgactatgga gagttcttgg cggctacgtt gcacatgaat aaactggaga gggaggaaaag 300
cttggtgtcg gcgtttgcgt tcttcgataa ggacgggagt ggcttcataa cgatcgacga 360
gctctcaciaa gcatgcgaac agttcgggtc ttctgatgtc catctcgagg atatgatcaa 420
agatgtggac caaaacaatg acggtcagat tgattatagt gagtttgccg cgatgatgag 480
aaagggcaat gctggcggag caggacggcg gaccatgagg aacagcttgc atgtgaatct 540
tggcgagctc ctgaagccca ccgagacctt gtttttttta accagaggat cctgcttccc 600
agatgcgccc catctgaaac ttaccgtgtg cctcatccgc tttcacagat cgattagtgt 660
atctctagta gcagcgccaa aagcagttta agattttctc agtggtacta gtatgtgagg 720
gaaaatagca tcctgtggtg ccagtggagag atttttaggcg ctgaagacga cgtgctatgc 780
tactccggag tcctgttg                                     798

```

&lt;210&gt; 37

&lt;211&gt; 632

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:723144.1:2000SEP08

&lt;400&gt; 37

```

attcctcctt cggatctcgc attcagcgca gcgcaggtgt gaaggaaggg ggacgcggag 60
gaagatgggg ctcgcgttcg ggaagctctt cagccggctc ttcgccaaga aggagatgcg 120
gatcctcatg gtcggcctcg acgcccgcg taaaaccacc atcctctaca agctcaagct 180
cggcgagatc gtcaccacca tccccacat cggtttcaat gttgaaactg ttgagtacaa 240
gaacattagc ttcactgtct gggatgtcgg gggtcaggac aagatcagac ctctttggag 300
gcattacttc cagaacaccc agggctcttat ttttgttggt gacagcaatg accgtgaccg 360
tgttggtgaa gccagagatg agctccacag gatgctgaac gaggatgagc tacgtgatgc 420
tgtgctgctt gtttttgcca acaagcaaga tcttcccaat gccatgaatg ctgctgagat 480
tactgacaag cttggattac actccctgcg ccagcgacac tggtacatcc agagcacttg 540
tgccacaact ggcgagggtc tgtatgaggg cctggactgg ctgtccagca acattgagag 600
caaggcctga ggcctacctt gaatatcaac tg                                     632

```

&lt;210&gt; 38

&lt;211&gt; 534

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1007188.1:2000SEP08

&lt;400&gt; 38

```

gaggaattct ggcgcttcta caagatgatg tccactcgcc gagacctcta cctgctcatg 60

```

```

ctgacctaca gcaaccataa ggaccacttg gatgcctccg acctgcagcg cttttctggag 120
gtggagcaga agatgagcgg cgtgacgctc gagagctgtc agagcatcat cgagcaattt 180
gagccttgtc tggagaataa gagcaaaggg gtgttgggga ttgatggctt tactaactac 240
acgcggagcc ccgccgtga catcttcaac cctgagcacc acagagtga ccaagacatg 300
acgcagcctc tgagtacta cttcatcacc tcctcccaca acacctacct cgtgggtgac 360
cagctcatgt cccagtctcg ggtggacatg tatgcctggg tcctgcaggg cggctgccgc 420
tgtgtggagg tggactgctg ggatgggcct gatggggaac ccattgtcca ccatggctac 480
actctgacct ccaagatcct tttcaaagat gtcacgaaa ccatcaacaa atat 534

```

&lt;210&gt; 39

&lt;211&gt; 436

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1024412.1:2000SEP08

&lt;400&gt; 39

```

gcgggaggca ggggtggaag ccttcctatt ggaacctcgg tgctgggaca gtgggactca 60
gagcagctct ctgacgcca ggtccacgct catggctgat gatgtcaggt actaacaacg 120
tcgcccaggg ccggaagctg gtggagcagt tgcgcacga agctgggacg gaacgcatca 180
aggctcccaa ggctcgtca gaactgatga gctactgtga gcaacatgcc cggaatgacc 240
ccttgctggt tgggtacca gcctctgaaa atccattcaa agacaaaaag ccttgcataa 300
ttctctagtt ctctctctct tcctccctcc ttctctctcc tctctttctc cccaccccaa 360
aggattgtc cagtttttcc tgtaagaaaa aaatatttga atcccttcta ctttgaattt 420
taaaggaaca gaaacc 436

```

&lt;210&gt; 40

&lt;211&gt; 2067

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:284797.3:2000SEP08

&lt;400&gt; 40

```

gagcggggag gaggatcgcc gggaccgggt ctgtgggggt tgagcgagag gtgagcgctg 60
ccgctcggcg catcattccg accgccctcc tgcgggctcg gaggaagccc cccgcgctgt 120
gcggaggcgc ctcggtgccc gggctgcgga gccccggccc agcaagaggg aaaagatgcc 180
gttccaccat gtgaccgccg gcttgttgta caagggggaa ttacctcaac cgatcgctct 240
ctgctggcag tgacagcgaa cagctggcta atatctccgt ggaggagctc gatgaaatcc 300
gagaggcctt tcgggttctg gaccgggatg ggaacggcct catctccaag caggagctgg 360
gcatggccat gcgctctttg ggtacatgc caagcgaggt ggagctggcc atcatcatgc 420
agcgcttggg catggacggg gatggccagg tggattttga tgaattcatg accattcttg 480
gccccaaact ggtgtcttca gaaggtcgcy atggttttct tgggaacacg atagacagca 540
tattctggca gtttgacatg caaaggataa ctctgggaaga gttgaagcac attctctatc 600
atgccttccg agaccaccta acgatgaagg acattgagaa catcattatc aatgaggaag 660
agagcctgaa tgagacctcg gggaaactgcc aaacagagtt tgaaggagtg cattcccaga 720
agcagaacag acagacctgc gtccggaaga gcctcatatg cgcttttget atggccttca 780
tcacagtggt catgctgatt gcagccaacc agatactccg gagcggcatg gtagtagcagc 840
ctcccgcag ccacactgtg ttgcagctca ccgcgcagtg gcatgccgtg cgggcagact 900
cttctccac acagcagaca cggacctatg gactatggat ggatgcggac gatggaaccg 960
ttcagtaaag aacccaaggt tgagtgaac tggggttgca ctccagcttc atctccttgg 1020
cagggacact aaagggctgt ctacagtgct ttaattggtt tgttttctaa tgcaataaat 1080
cgccaggatt tccaaattat tgcttcaacc atcagtcata actctagaga aaaacttcat 1140
cttcatctcg gaattccagc agccatccag gtataatggt gagagagcaa gagagaggca 1200
tgaacacaca gtcacctcca ctccccattc tttcaaccca aggggtcaaca ggagactctc 1260
tgctgaggtc aacctcagt tccaagaggg tgacttcagg gttgccacag aagagggggc 1320
tttccatccc accctggctt ccttctccca ccagtgagg agccccccaa ggcatgagg 1380
gaaagccatt aagtggggtg cagcccctgg cggtcttcag ctgcacccct caccgcttcc 1440
gttcttactg ccttttctag aggactgcca ttggtagagc tggggtgctc tcagtttccg 1500

```

```

tcggcacttg gctttctgtt cctagagccc atccccgacc agcatttttg aacctgcaga 1560
gagcctttct cccagctttg ccgcctgtgc tgccattaga gaaaaaaga agcataatcg 1620
tgggcatctc tataacatag acagatatat atatttggct tcttcgtgtg tgatgtcatt 1680
cttccctaata cgcttacacc taatataagc ttgtcctccc tcccacactc atccctgctt 1740
ctttggggca ggggtgaggg aggttgaggg gggaggggac tgggactggc ggtgaggttt 1800
gatacccaca cccagccatc tggagtcagt gagaccacc tctgtctctat gccggatgac 1860
cttghtaagtt aaggtccttg tgcattcttac ctttccctta ccttggagtt tgcttgttag 1920
tcttccccag ctaaaaagag aaatgcaatc caaaatggaa agatggagag atcaaacgac 1980
tgtattttca cataagatct tctggccttc agattctgca actgactcat ttcaggaggt 2040
tgggggaaaa gatgtttatg agtttagc
2067

```

&lt;210&gt; 41

&lt;211&gt; 640

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1092901.1:2000SEP08

&lt;400&gt; 41

```

gaagagtacc agaagagtct gccagagcag taccatctgg gtctggatca aaaacgcaga 60
aaatacgtgg ttggagagct tatcggaat tctgccgatt tcacgactat gcagtgaggt 120
gctggggctg atggtgatga tctgaagcat gctcaggaga gaggtggtgc agatcaagag 180
gcccctcatt accttgtata tgtagaacac agccttacat ctgaagttat tctgaaaatt 240
ctgtgacgca aacacgtcta gagacaataa ctctgctgcc atgactagca tcaactatgta 300
gacagaaacc aagtgcagca tgatcaggtc atgggtttttt aggccttgta tccttaaaga 360
atgtaaaaat gtgccagaga agaagaaaga tggtggctga gattccaatg ccagcttgaa 420
aacaaaaggc cttttgaaag ataacatgtg aaatgtgtgt tcactcctaat gacatagaag 480
aaacattttg aagacctgaa aaaagtagat ctgtgtcatc aatgtccttt tgtcagtgtt 540
caaaattatt accattatta ttcttatttt actctcattt cttgattaac cccatcttat 600
ataaattctt gatcatgtat ctctatact atttctctaaa
640

```

&lt;210&gt; 42

&lt;211&gt; 752

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:228930.1:2000SEP08

&lt;400&gt; 42

```

ccccagtccc tgccgcgacg gcccgggccg cccggagccc agatgagccc agatggccgg 60
ggctcaaccc ggagtgcacg ccttgcaact caagcccgtg tgcgtgtccg acagcctcaa 120
gaagggcacc aaattcgta agtgggatga tgactcaact attgttactc caattatttt 180
gaggactgac cctcagggat tttactttta caggacagat caaaacaagg agacagagct 240
actggatctc agccttgtca aagatgccag atgtgggaga caccgcaaag ctcccaagga 300
ccccaaatta cgtgaacttt tggatgtagg gaacatcggg cgccctggagc agcgcatgat 360
cacagtgggtg tatgggcctg acctcgtgaa catctcccat ttgaatctcg tggctttcca 420
agaagaagtg gccaaaggaat ggacaaatga ggttttcagt ttggcaacaa acctgctggc 480
ccaaaacatg tccagggatg catctctgga aaaagcctat actaaactta agctgcaagt 540
cactccagaa gggcgatttc ctctcaaaaa catatatcgc ttgttttcag cagatcgga 600
gcgagttgaa actgcttttag aggccttgtag tcttccatct tcaaggggtg aaaaagctaa 660
tgaggctgag aaaaagtgagc agagctgtgg aaaggctccc ccaaaacatt ttcaccttca 720
ttttataaag caaaataaaa tgcttgaaac tg
752

```

&lt;210&gt; 43

&lt;211&gt; 928

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;



&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:722913.1:2000SEP08

&lt;400&gt; 43

```
atccacacac atccgcaccc ctctctctctc gggtttctctt cccaccacca acccccagat 60
cgcagcgatc tccgcgcgcg cctctctctg gattccgcgc gacaaagtca ggcgtagagg 120
ctccaggagg aggagggagg cgcagcaggg cgggtggggg agatgttcct ctgggactgg 180
ttctacgggg tgctggcctc cctcggcctg tggcagaagg aggccaaagt cctcttcctt 240
ggcctcgaca acgcccggcaa gaccacgctg ctccacatgc tcaaggacga gcggttggtg 300
cagcaccagc cgacgcagca cccgacgctg gaggagctca gcatcggcaa gatcaagttc 360
aaggcggttcg acctcggcgg ccaccagatc gcgcgcgcgc tgtggaagga ttactacgca 420
aaggttgatg ctgtagtata cctggttagat gcatatgata aggagcgatt tgctgaatca 480
aagaaggagc tcgatgctct cctgtctgat gattcttttg ccaatgttcc atttctcctc 540
cttggcaaca agattgatat cccatatgct gcctctgaag aggagctacg gtatcaccta 600
ggccttagca acttcacaac cggaaggggc aaggtcaacc ttggcgactc caatgtccgt 660
ccacttgagg tcttcatgtg cagtgttgtt cgcaagatgg gctacggtga tggtttcaag 720
tggtgtctccc agtacatcaa gtagggccct cgccgaggcg catgctcagc ccggcgaatt 780
tttttgcctt gttcgttccg gttggttcgt gacgagaacg aacgatggtc gtctaggaaa 840
gtatcatgga gggatgaag aggaaaaaca gtgcttaaca ttgttgaaca tgttgagtct 900
ttagcgctcaa aacaatttat gtcgaact 928
```

&lt;210&gt; 44

&lt;211&gt; 561

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:457478.1:2000SEP08

&lt;400&gt; 44

```
gccggcgggg cgcgcggcgg tgccggcggg tgactggcgg cgggcgcgcg ggtcgggctg 60
gctgccgggc agcatggagg agctgagcag cgtgggcgag caggtcttcg ccgccgagtg 120
catcctgagc aagcggctcc gcaagggcaa gctggagtag ctggtcaagt ggcgcggctg 180
gtcctccaaa cataacagct gggagccgga ggagaacatc ctggaccgga ggtgctcctt 240
ggccttccag aagaagttag gacgctgaca gcaactggga ggggtgtggg gagggacggt 300
ggctgggttg agtcgtgctg ggcgctgctc gcctgccggg aggctggagc tggggctgtg 360
gtctgaggtc ccttagcaag attgtgtcca gggccatctc tgtggctggc tgtaaaagtc 420
cccagccagc cctctcccca ccttctccta gccagcctgt atccatccct gagaagaacc 480
taagtaagtg ctgcttgggt gtttccgaca gccatacttg caaaaaccaa ctttttcagt 540
tgcttcagat aaagatgtta t 561
```

&lt;210&gt; 45

&lt;211&gt; 1247

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:358719.1:2000SEP08

&lt;400&gt; 45

```
gatttggcca ttttcgcggg aaaactgaat aagaggaagt gaaatctgaa taattttgtg 60
ttactcatag cgcgtaatat ttgtctaggg ccgcgggggac tttgaccgtt tacgtggaga 120
ctcgcocagg tgtttttctc aggtgttttc cgcgttccgg gtcaaagttg gcgtttttatt 180
attatagtca gctgacgtgt agtgatttta tacccggtga gttoctcaag aggccactct 240
tgagtgccag cgagtagagt tttctcctcc gagccgctcc gacaccggga ctgaaaatga 300
gacatattat ctgccacgga ggtgttatta ccgaagaaat ggccgcccagt cttttggacc 360
agctgatcga agaggtagtg gctgataatc ttccacctcc tagccatttt gaaccaccta 420
cccttcacga actgtatgat ttagacgtga cgcccccgga agatcccaac gaggaggcgg 480
tttcgcagat ttttccgac tctgtaatgt tggcggtgca ggaagggatt gacttactca 540
cttttccgcc ggcgcggcgt tctccggagc cgcctcacct ttcccggcag cccgagcagc 600
cggagcagag agccttgggt ccggtttcta tgccaaacct tgtaccggag gtgatcagtc 660
```

```

ttacctgccg cgaggetggc tttccaccca gtgacgacga ggatgaagag ggtgaggagt 720
ttgtgttaga ttatgtggag cccccgggc acggttgacg gtcttgtcat tatcaccgga 780
ggaatacggg ggaccagat attatgtgtt cgctttgcta tatgaggacc tgtggcatgt 840
ttgtctacag tcctgtgtct gaacctgagc ctgagcccgga gccagaaccg gagcctgcaa 900
gacctacccg ccgtcctaaa atggcgccctg ctatcctgag acgcccgcaca tcacctgtgt 960
ctagagaatg caatagtagt acggatagct gtgactccgg tcctttctaac acacctcctg 1020
agatacaccg ggtgggtccc ctgtgcccc ttaaaccagt tgccgtgaga gttgggtggg 1080
gtcgccaggc tgtggaatgt atcgaggact tgcttaacga gcctgggcaa cctttggact 1140
tgagctgtaa acgccccagg ccataaggtg taaacctgtg attgcgtgtg tgggttaacg 1200
ctttgtttgc tgaatgagtt gatgtaagtt taataaaggg tgagatc 1247

```

&lt;210&gt; 46

&lt;211&gt; 1017

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:105160.5:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 977-978

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 46

```

gccacagagc gccaccatcg agatgtcgta gagttactta tcaaatatgg agctgatgtc 60
catgctttca gcaaatltga taaatcagcc tttgacatag ctctggagaa aaacaatgct 120
gagattttgg tcacctcca ggaagcaatg cagaatcagg tgaatgttaa tcagagaga 180
gccaaccctg tgactgaccc tgtgagtatg gctgctccat tcatcttcac gtcgggtgag 240
gttggttaacc tcgcaagcct tattttctca accaacacca aaacaacctc agccaatata 300
gaggaaatta tagaaggaaa ttccgttgac tcatcaatcc agcaagtaat ggggagtggg 360
ggccagaggg tcatacccat agtgactgat ggagtccctc tgggtaatat ccaaacttca 420
atccctactg gaggcattgg ccagccattt attgtaactg tgcaagatgg acagcaagga 480
aggcaatgaa agagagctac tacagcaaca actccaggag gccaatcgaa gagcccagga 540
ataccgacac cagctcctaa agaaagagca ggaagcagaa cagtaccgtc ttaagctgga 600
ggccatagcc cgacagcagc ccaatggagt tgatttcacc atggttgaag aggtggctga 660
ggtagatgct gtagtagtca cagaggggga gttggaagag agagagacaa aagtgactgg 720
gtcagcaggg accacagagc ctacacttag agtttccatg gcaactgttt catcttaata 780
tgcaagggcc acaatttgca ctgtgttcat attaatcctc ttttaaaaaa ggaaatatac 840
agaagacaaa cattgtataa aaactaagag tgtctttaag aagaaaacta tagcagggta 900
caatgcttgg gtcaggaag tttctctgtg caactagaaa attcaaagcc atatttaggg 960
aacatttttt ctgaggnncc aaaagaataa aggaccaa atcttagctc atatcat 1017

```

&lt;210&gt; 47

&lt;211&gt; 1004

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:400705.1:2000SEP08

&lt;400&gt; 47

```

ggaggcggca tgagcagcgc ggcacagagc tgacgccgcg ccccgccggg ccccatgtcc 60
ttcgccacgc tgcgcccggc gccgcccggc cgctacctgt acccgagggt gagcccgtg 120
tcggaggagc aggaccgagg cagcgacagc tcggggtccg acgagaaacc ctgtcgctg 180
cacgcggcgc gctgcggcct ccagggcgcc cgcgggaggg cggggggccg gcggggcggg 240
ggcggggggc cagggggggc gccaggccgt gagccccggc agcggcacac ggcgaacgcg 300
cgcgagcgag accgcaccaa cagcgtgaac acggccttca cggcgctgcg cagcgtgac 360
cccaccgagc ccgcccagcg caagctctcc aagattgaga cgctgcgcct ggccttcagc 420
tacatctcgc acctgggcaa cgtgctgctg gcggggcagg cctgcgggac ggacagccct 480
gccactccgg gccgccttc tccacgcgg cgcgcgccg cagcccccg ccgcccggc 540

```

```
cgccgcctcc cgcccgcgac ggcgagaaca cccagcccaa acagatctgc accttctgcc 600
tcagcaacca gagaaagtgt agcaaggacc ggcagagaaa gacagcgatt cgcagttagg 660
aggtggccgg cagcagccag gaggcagacg ctgctggggg aggtggacgc ccggggtgac 720
tgcagacagc cccacacctg gacctgagct gggcaaggcc caccgcaagc atgccccag 780
gccagccctg gctgcgagcg gggccgaggg acagacggac gtacagacag gcgcccgcag 840
cgggactctg cgctggcccc agcacctgcc cgggcccact ggaactttct gcgctggctt 900
ttcttccggc cactgtgtgt atggcatctt gtgtttttga tatgataata taaagtctga 960
aaattttgta taattaaaaa caaacagta tcttccaaaa aaaa 1004
```

&lt;210&gt; 48

&lt;211&gt; 1642

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:221977.1:2000SEP08

&lt;400&gt; 48

```
ggagcacaag ggcgagctct cctccgccct gcagcctcag gtaaggatcc ggaccttcac 60
tgctgcccc aactggggctc ctccagctcgg ggcctctagg ggcttgggct gctgggtctgg 120
gtggaggcgc gttgggggtg tcgaggagct tgggtgcaag ccccatgcgg gaagctcagg 180
gcaaagcgcc ttgaaacctt ctctagttaa ggaaggtaaa atggctccgg gaccccaggc 240
ggggaagag gtgtccaagg aggagggctg gagagtaagg cggggatggc agagggtagc 300
cagaggctga gaagtgggac cagtgtctca agggcgccca gcctcttctt ccgccaaca 360
tggccaccat ttttgttggg atagccttgg gctcctgcc a ggggaacagc agggcagtga 420
gggagtcagg actttgggct ggtggggagg ccgagaatca aggtctctgg gagtcaggac 480
tttggggctg gtggagtggc tggggactga ggccttgggg ggtcaggacc ttgggctgct 540
atggaggctg ggactgaggc tctgggggag tcaggacctt gagctggtgg ggaggctggg 600
aattgaggcc ctggtgtgtt caagaccttg ggtctggtgg gaggccaggg cctgaggccc 660
tggggggtca ggaccttggg ctgttgggga ggctggggac ttagggcctg gggggtcagg 720
atcttgggct ggtggggagg ctgaagggtg gggactgagg ctctgggcag tcaggacctc 780
agggttgggg tggactggag accaagacct tggcagggga gctgtggctg cctcatttgg 840
acgctggagg gtggctagcg tggctggaag cggccagcca ggggttcaca gaactgaagg 900
tgaggcctcc agaggcccta gtctctacct gactgtctct gaaactgggg ggatgggggtg 960
gagcctttag ggggaaggga agagggaact gaagaggaag tgggggaggg aggtagagga 1020
ggcaagtctg gcgccatgct gactcaccgc ccacaaggcc cagggcgggc cctcgggggg 1080
ccttggcagg gttgggggga tcttaggaag ccacaaggag ggctgggggg ctcttggagc 1140
aggagttagg aggcctgggc agcctgaaga gtacacgccg acggacagac agacagtga 1200
gtcaccata aagtagaaag cactactaac agcactggag ggtgtagtgt ttctacttt 1260
atggatgagt gtactgtggg cttcggagat cagccactg ctgccgcccg ctgcccgcca 1320
ccatcttctt cggcgctcgg ggacctcgtg tgacaggtga gcaccttacg gcccctcctt 1380
acctgcccc gatgcctgaa aggcctccat ggctttcctg cccttccctg ttccggacag 1440
ctggggaaag gccacagcag ctccctctgct gccctgcagt ctttgggggc ggggagggtt 1500
ggacatgtgg aacctgatg cagccgcagc gtcaaggacg aggaaggggt gggaagggtt 1560
ggtacgtgga ggggaatggg tgggtgggac agggacccaa tgctaataaa gactggactg 1620
tgcttcaaaa aaaaaaaaaa gg 1642
```

&lt;210&gt; 49

&lt;211&gt; 2303

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:898771.1:2000SEP08

&lt;400&gt; 49

```
tcggtagcgc agagcccggg gccggcgctg gcgcggtcgg agggccgcga gttcagagttc 60
ctcatgcgcc agcccagcgt caccatcgcc cgcaactcgt cgcagggctc ggtggacttg 120
agcatgggcc tgtccagctt catctcgccg cgccacctgc agctcagctt ccaggagcgc 180
cacttctacc tgcgctgcct cggcaagaac ggcgtcttcg tggacggggc cttccagaga 240
cgcgccgcgc ccgccctgca gctgcccagg cagtgtacct tccggtttcc cagcacggcc 300
```

```
atcaagatcc agttcaacttc gctctatcac aaagaagagg cccagcctc cccgctgcgg 360
ccactgtacc ccagatctc ccctctgaag atccacatcc cggagccgga cctccggagc 420
atggtcagcc ccgtoocctc ccgcagggc accatcagtg tccccaactc ctgccagcc 480
agtccacgcg gtgcgggctc ctccagttac cgctttgtgc agaactgtac ctccgacctg 540
cagctggcag cagagtttgc agcaaaggcc gcgtcggagc agcaggcaga cactgtctgga 600
ggagacagcc ccaaggtctg agcccacctg gcgccgtggt gcacctggtg acccaggatc 660
gggccatagt gaccggcgcc gggcggcaca gacagcccag ggccctgcggg cgtgtgctcg 720
agggtgtctc agcccacagg gaatggagcc gcagctggcg gtagagccgt cggcccttcc 780
tcgttttctg gcccaaatca tcccgttgg tcatttgctg gtgacagggg tggcgccgtc 840
accgagggcc cctcccgcca cgctccttgt tcatctgccg ctggcagccc ggcgtcaaag 900
ccttcataca gcccttgagc aggatggcgt tctccttata actttgtctt agttcatttt 960
tgtccaatag ctgtttttcaa attgatttgt ggaagagggtg agaaatttca gaattaaatt 1020
ataataaaac cgaagttttc ctcccatga ggcaaaaatg gttgttccta gttgaatgca 1080
agcacattaa tgggatagtt ggagggcact ggagcgggtg ccggcaagac gcgcacattc 1140
tcgtgggaag gttcctcccc atcagctgtg acgggtgggt ggggtggact cagggacctt 1200
gctggagctc atcccgtgag ccgtgtttct ccgcaggatg agtcaaagcc gccgttctcc 1260
tacgcgcagc tgatcgtgca ggcacatctc tccgccagc accggcagct gacctgagc 1320
gggatctacg cccacatcac caagcattac ccctactacc ggacggccga caaaggctgg 1380
caggtgaagc cgagtcacca gggccggatc gcctctgaag gcccttagaa catggaaactc 1440
acaggcatat tgcttccctt aaaacagaag agggccagtg agggggctca cgcttgaac 1500
cctagcattt tgtgaggccg aggcaggagg aggatcaaga tccagttcac gtcgctctat 1560
cacaagaag agggcccagc ctcccgcgtg cggccactgt acccccagat ctcccctctg 1620
aagatccaca tcccggagcc ggacctccgg agcatggtca gcccctgcc ctcccgcagc 1680
ggcaccatca gtgtcccaa ctctgccca gccagtccac gcggtgccgg ctctccagt 1740
taccgctttg tgcagaacgt gacctcggac ctgcagctgg cagcagagtt tgcagcaag 1800
gccgcgtcgg agcagcagcc agacacgtct ggaggagaca gcccgaagga tgagtcaaag 1860
ccgcgcttct cctacgcgca gctgatcgtg caggccatct cctccgcccc ggaccggcag 1920
ctgacctga gcgggatcta cgcccacatc accaagcatt acccctacta ccggacggcc 1980
gacaaaggct ggcagaattc tatccggcac aacctctctt tgaaccgtta ctttatcaaa 2040
gtcccacgtt ccaggagga gcctgggaag gggctccttt ggccaataga ccctgcctct 2100
gaagccaagc tcgtggaaca ggcattccgg aaacggaggc agaggggtgt ctctgtcttc 2160
cgcacccctc tcgggcctct gtctcaagg agcgctccag ctccgcccac acaccccggy 2220
ctgatgtccc ctgcgtccgg cggcctgcag accccagagt gcctgtctcg ggagggctcc 2280
cccattccac acaacatacg agc 2303
```

&lt;210&gt; 50

&lt;211&gt; 571

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:457478.1:2000SEP08

&lt;400&gt; 50

```
gccggcgggg cgcgcgggcg tgcggggcgg tgactggcgg cgggcgcgcg ggtcgggctg 60
gctgccgggg agcatggagg agctgagcag cgtgggagcg caggtcttcg ccgccgagtg 120
catcctgagc aagcgggtcc gcaagggcaa gctggagtac ctgggtcaagt ggcgcggtg 180
gtcctccaaa cataacagct gggagccgga ggagaacatc ctggaccgga ggctgtctct 240
ggccttccag aagaagtgag gacgctgaca gcaactgtgg gagggtgtgg gggagggacg 300
gtggctgggt ggagtcgtgc tgggcgctgc tcgcctgccg ggaggctgga gctggggctg 360
tggtctgagg tcccttagca agattgtgtc cagggccatc tctgtggctg gctgtaaaag 420
tccccagcca gccctctctc cacttctctc tagccagcct gtatccatcc ctgagaagaa 480
cctaagtaag tgctgctttg tgttttccga cagccatact tgcaaaaacc aactttttca 540
gttgcttcag ataaagatgt tatagcttca a 571
```

&lt;210&gt; 51

&lt;211&gt; 1410

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: LI:125140.1:2000SEP08

<400> 51

```
ccctccccgc cggggcgcgcg tgcccgcgagc tccacgcgcg ccaggcccgcc ctggggaggg 60
cgcgcgcgcg agcgggcgacc ggagaggagc tcggggcggc ggaggtgatg acggcgggcg 120
catggagttc cctgagcacg gcggacggct gctggggcgcg ctgaggcagc agcgcgagct 180
gggcttcccta tgcgactgca ccgtgctggt gggcgacgcg cgcttcccgg cccaccgtgc 240
cgtgctggcc gcgtgcagcg tctacttcca tctctttctac agggaccggc ccgcgggcag 300
tcgcgacacg gtgcggctca acggcgacat cgtcacggcg cccgccttcg gccgcctact 360
ggacttcatg tacgagggcc gcctggacct ggcgacgcct gcctgtggag gacgtccctg 420
cagccgccag ctacctgcac atgtatgaca tcgtcaaggc ctgcaagggc aggctccagg 480
agaaggatcg aagtctggac ccggggaacc ctgcccctgg ggagaaacct gctcagccac 540
cgtgcccctg gcctgtctgg accgcgacc tctgcccagc tgcccgaag gccaaagctcc 600
ccccgtttgg ggtcaaggct gccctccctc ctcgagcatc tgggcctcct ccctgccagg 660
tcccagaaga gtcagaccag gccctggacc tatcgttgaa gtctggccca aggcaggagc 720
gggtccacc cccgtgcgtc ctccagacac ccctctgcag ccagaggcag ccagggggcc 780
agccactggt gaaggacgaa cgggactcac tgtccaaaca ggaggagatc agcagctcta 840
ggagccccca cagtcccccg aagccacctc ctgttcctgc agccaagggc ctggtggtgg 900
gcttgacgac gctgcccctc agcggagagg gcagccggga gctggagctt ggtgcagggc 960
gactggcgag tgaggacgag ctggggcctg gtggggccct ctgcatctgc ccgttgtgca 1020
gcaagctggt tcccagctcc cacgtgctgc agctgcacct cagtggccac ttccgtgagc 1080
gagacagcac ccgggcccgg atctcaccgc acggcggtgg acccacctgc ccgctctgtg 1140
ggaagacctt ctggtgcaca tacacactga agaggcacga gcggacacac tcgggtgaga 1200
agccctatac gtgtgtgcag tgtggcaaaa gttttcagta ctcccacaac ctgagccggc 1260
acaccgtagt gcacactcga gagaagccgc atgcctgccg gtggtgtgag cgccgtttca 1320
cgcagtcgag ggacctctac cgccacgtcc gcaagtttca ctgtggcctc gtcaagctcc 1380
ttctggtgtg atgcatccct gtggtcctga 1410
```

<210> 52

<211> 271

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:021095.2:2000SEP08

<220>

<221> unsure

<222> 23

<223> a, t, c, g, or other

<400> 52

```
tgtcttggtg cgggcaccga gtnccgctg aagatggcgt ctaccagccg tttggatgct 60
cttccaagag tcacatgtcc aaaccatcca gatgcgattt tagtggagga ctacagagcc 120
ggtgatatga tctgtcctga atgtgggctg gttgtagggc acaggagctg caagttttga 180
cgaatttggc aattctaagt accagaatcg gagaacaatg agcagttctg atcgggcaat 240
gatgaatgca tccaagaaat cactaccatg g 271
```

<210> 53

<211> 921

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:888730.1:2000SEP08

<400> 53

```
cccaggatgc tgtatcttgc ttccaacagt gtttggacgg aacagaccgc gggactccca 60
ctttattcac ttccagccgc cttacaagtc ttcttccaag tcgcagcctc cgggaccatc 120
tccttgccgc cttcggtctc taggaccagc cagccccgtc ttccgggtta gctccatact 180
ccggatcagc catgacctct cagattcgtc agaattattc caccgaagtg gaagctgccg 240
```

```

tgaaccgcct ggtcaacttg cacctgcggg cctcttacac ctacctctct ctgggcttct 300
ttttggatcg ggatgacgtg gctttggaag gcgtaggcca cttcttaccg cgaattggcc 360
gaggagaagc gcgagggcgc cgagcgtctc ctcaagttgc agaacgaacg cgggggcccgt 420
gcaactcttc aggatgtgca gaagccatct caagatgagt ggggtaaaac cctggaggcc 480
atggaagctg ccttggccct ggagaagaac ctgaaccagg ccctcttgga tctgcacgcc 540
ctgggctctg cccgcacaga cctcacctc tgtgacttct tggaaagcca ctccctggat 600
aaggaggtga agctcatgca tagaagatgg gcaaccacct ggacctaac tccgttaggg 660
tggcaggggc acaaccacgc gtcagactgg cgtggcccag gcatactctg ggcagatc 720
tcatttgagc gcctcacatc tgaagcacga ctaggaggcc tctgtacctt cccaaggggc 780
tctctcctc tgctctgcac cagcccgccc tgggacctcc acctgaatga acctctcaag 840
ccactaggca gctttgtaac cgccctggag cctctgtcaa gtcttgacc aagtaaaaat 900
aaagcttttt gagacagcaa a 921

```

&lt;210&gt; 54

&lt;211&gt; 1315

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:358719.1:2000SEP08

&lt;400&gt; 54

```

gatttggcca ttttcgcggg aaaactgaat aagaggaagt gaaatctgaa taattttgtg 60
ttactcatag cgcgtaatat ttgtctaggg ccgcggggac tttgaccgtt tacgttggag 120
actcgcccca caggtgtttt tctcaggtgt tttccgcgtt ccgggtcaaa gttggcggtt 180
tattattata gtcagctgac gtgtagtgta ttataaccg gtgagttcct caagaggcca 240
ctcttgagtg ccagcgagta gagttttctc ctccgagctc gctcgcgaca ccggcacctg 300
aaaatgagac aatatgtatt ctgccacgga ggggtgttatt accgaagaaa tggaaacgaca 360
gatcttttgg accagctgat cgaagaggta ctggctgata atgttcccac ctccatagcc 420
attttgaacc acctaccctt gcacgaactg atatgattta agacgtgacg gccccgaag 480
atcccaaaga agaagccggt ttccagaatt ttcccggat cctgaaatgt tggcggtgca 540
ggaaggggatt gacttactca cttttccgcc ggcgccccgg tatctccgga gccgccgtca 600
cgctttcgcc ggcagcccca gcagccggag ccagagagcc ttgggtcccg gttttctatg 660
gccaaaaccc ttgtaccgga ggtgtatcga tctttacctg ccacgaggct ggctttccac 720
ccacgtgacg acgaggattg aagaggggtga ggagtttgtg tatagattat gtggagcacc 780
ccgggcacgg ttgcaggtct tgtcattatc accggaggaa tacgggggac ccagatatta 840
tgtgttcgct ttgctatatg aggacctgtg gcattgtgcg tctacagtcc tgtgtctgaa 900
cctgaggcct gagcccgagc cagaacctgg agcctgcaag acctaccgcg cgtcctaaaa 960
tggcgctgct tatcctgaga cgcccgga caacctgtgt ctagagaatg caatagtagt 1020
acgcatagct gtgactccgg tctttctaac acacctgcct gagatacaac ccggtgggtcc 1080
cgctgtgccc cattaaacca gttgcogtga gagttgggtg gcgtcgccag gctttgggaa 1140
tggatcgagg accttgcttaa cgagcctggg caaccttttg acctgagctg taaaagcccc 1200
aggccataag gtgtaaactg gaaaatagcg tgtgtgggtta acgcctttgt ttgctgaatg 1260
agttgatgta agtttaataa aggtgagat aatgttaaaa aaaaagaaaa aaagg 1315

```

&lt;210&gt; 55

&lt;211&gt; 1484

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:351342.3:2000SEP08

&lt;400&gt; 55

```

accgtctgga tgactccacc agccgcgggg attgtggcaa acgcaatggt gggtttctcc 60
tccaggcctc tggcctcgct gccaggtcc agcacggtg cgcccgccgc atcatgggag 120
ccccccgccc cgccctggct ggtgaggatg tatccgttgg ccgagttggc agatgtgggt 180
accaccctga ccatgggtgac ggtggggggc tgctgcacga cgtggatggc gtggcccgcc 240
ggctgctgtg aggttacgat ggaggcgggc atgtaggcca cgggcttggc cagaggctg 300
gacggtcggg gaggcacggc catgatcact ggctggggcg tgacggggga gccgggtgcg 360
ctttgggaat accggtactc tgggacagaa gctaacttgg acccaaaactc agggctcgtg 420

```

```

ggaatggggg agccctcccc agacaggcac tctgggggtct gcaggccgcc ggagcgaggg 480
gacatcagcc cggggtgtgt gggcgaagct ggagcgctcc ttgaggacag aggcccgaa 540
ggggtgcgga agcaggagac accctcttgc ctccgtttcc ggaatgcctg ttccacgagc 600
ttggcttcag aggcagggtc tattcgccaa aaggaccctt tcccaggctc ctctctggaa 660
cgtgggactt tgataaagta acggttcaaa gagaggttgt gccggataga attctgccag 720
cctttgtcgg ccgtccggta gtaggggtaa tgcttgggtg tgtgggcgta gatcccgctc 780
agggtcagct gccggtcctg ggcggaggag atggcctgca cgatcagctg cgcgtaggag 840
aacggcggtt ttgactcatc cttggggctg tctcctccag acgtgtctgc ctgctgctcc 900
gacggcggtt ttgctgcaaa ctctgctgcc agctgcaggt ccgagggtcac gttctgcaca 960
aagcggtaac tggaggagcc ggcaccgctg ggactggctg ggcaggagtt ggggacactg 1020
atggtgcccc tcggggaggg gacggggctg accatgctcc ggagggtccg ctccgggatg 1080
tggatcttca gaggggagat ctgggggtac agtggccgca gcggggaggc tggggcctct 1140
tctttgtgat agagcgacgt gaactggatc ttgatggcgg tgctgggaaa ccggaaggta 1200
cactgcttgg gcagctgcag ggcggggcgg ccgctctctt ggaaggcccc gtccacgaag 1260
acgcccgttct tgccgaggca gcgcaggtag aagtgcggct cctggaagct gagctgcagg 1320
tgccgcccgg agatgaagct ggacaggccc atgctcaagt ccaccgagcc ctgcgacgag 1380
ttgcggccga tggtagcgt gggctggcgc atgaggaact cgaactcgcg gccctccagc 1440
cgcgcagcgg ccggccccgg gctctgccgt accgaggcgg ccgc 1484

```

&lt;210&gt; 56

&lt;211&gt; 1996

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:256099.2:2000SEP08

&lt;400&gt; 56

```

cctcccgccg cggcgcgcat ctcatccag ccctcattcc gcgcattcca gcgtcctcct 60
cgcacactcg aggccagggg gcgggagggc gcgagctccg gcgcgccgcg gtccccccag 120
aacgatcgcc gcggccccga agagttggcg ctccggggcg actccttgga actgtgatta 180
gcgcacccat cccaccttcc cggaccctgg gaccgggtcca acgagcgctc ctccaaagcg 240
gagcgcttgg agggcacagg tcggctacgc attaccatcc aacgtatgtg gagctctccc 300
acctcccttg taggggagca ccggctctgg gggcagtcag gctactggac aagccccatcg 360
acaatggtac gcagagggcg tgtggagccc ggatattgta gcagagtttc caggaggccc 420
tcgggccatc tacccgccct gtgtgcatgt gctgcaaaat catcctgtct ggacgagggc 480
aagatgtatg ggtcggaacg atgctgattg ccgctacat caagtctccg gacagggaag 540
atccccgtcaa ccaggaaggc aggtctccag ccacatccca ggtgctaggc tcgtcgcata 600
agctcgctga gatccatgtg ccaatgctaa atggaccagg cagctaagga caaggcccg 660
gcgagaagac atgtgctagc gcagtgctgt ctgcacagat ccatctccgc cacaggcctt 720
tcacagtaaa gcataggccc tcgccccggg gccccgggtc tgcccagcta gtactcgagg 780
gtttgcggca agggagcttg tcgcaggggc caaagcccgg aacgtcccca tgatgtgaaa 840
gccttgtctc tcagcgaaat ctatgcgtgt ccagtcctcc gtttgcctct gcctagggtt 900
ttgagtctcc tggcagggcc cgtgcceat cgccctctgg cgccccagg caccacacca 960
tggcagggca cacaagcagt gggccaagct cacaagcctc gtggatgttg gagttcttca 1020
tgcacattca catggagcaa gcaagcaagg acaccaggac acgttacaac aagacacctg 1080
ttcagtgcac attggccatg tccagcccca agcgtatacg agagaccctt atatctacag 1140
acgccgtgca ctactccgcc ataattctat gacaaaattc accaggtaga taaaatgggt 1200
ggacacagag gatcatcttc agacaggggt acaccatcca aatgcctttg tttcttgg 1260
aagttctggg caagacactc aacaccatac atcgaggatg caaggcagct ccttctatgt 1320
gggtctccag cccagtatgt agagcccccg aggaacatgt atcattacct gctccacgaa 1380
gtgtctgctt ctttcgggca agccagggtg gtggcagaaa agttgacgac acgagtatgc 1440
tcgctatgta gatatggaca ctacgtctta ccggcatcca ccgggtcccc gctcttgtga 1500
gtacagtgat caacttcate cagcaagctc aagcacctcc cgtgagaagt aacctagt 1560
gtaacacgcg ctgcctggag gaaacttcta caccacttcc tcgcacaggt gggtctcacc 1620
acttcacgat gagacaccac acacggatga tcctttgctg tgcacattgt cctaattgtct 1680
ttgatggtgt ctatcttcca cgtgaggcac cggggctcag caccacactc ttacaggctg 1740
tgtgtaagta ctgacgagac tctggggagg cagggtatggg cgggcacgag acggtgtgct 1800
gccaggatac tcgcggaccg tctcggcacg gggaccccg cagggccgcc acggcacc 1860
tgaagtgcac caagagaagc ctgagaagga gcacagattg tgactctacc cagggaccaa 1920
ctgtgcctga acctgaggtg ccccaacccc aataaacccc agatgctgtg tattttccaa 1980
aacacaaaaa aaaagg 1996

```

<210> 57  
<211> 1438  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:2051991.1:2000SEP08

<400> 57  
ggttcactgt ggttgatca tgcacatgaa aggaaagggc agatgaaagc tgggaaggagc 60  
acaagccaag cacagtggga ggggcacatg cagggggctg catggtggca tgcagggttc 120  
ccacggagct ttcagtttgc actcagggct gtgaggtcat ggaggccagc attgcacttc 180  
tcatggcagg tgtccacgga gtcacatgaa tcatgtgggt ttcaccacac agccagcacc 240  
tttgcctgca acctctgagt ttctgttag cagtttttgg gttgctgtga tgaatgagac 300  
catatccgta atatcacagc gtgtatttct gtctttccaa ggatgtgtga ccggaatggt 360  
ggtcggcggc ttcgacagt gctgatcgag cagattgaca gtagcatgta tccaggactg 420  
atgtgggaga atgaggagaa gagcatgttc cggatccctt gggaacacgc tggccagcca 480  
gattataatc aggaagctgg atgcctccat ttttaaggcc tgggcagttt ttaaagggaa 540  
gtttaaagaa ggggaccaag gttgaaccag ccaacttgaa gacgcagggt acgcctgatg 600  
catttgaaata acgagcccg tattttgagg aacgtgatcg gaccgggttc caatctggga 660  
acattttccg atgccatc caagtttacc gaattcgttc ctgaggaaga gcaaagaact 720  
gtaactagtc ctttgatggg catgaatacc ttccacgaag ctgdcctggt cctgtagacc 780  
ttcgaccaca gaacttgggt ctggggtaga cacagggtct ggggcacata ttttcagttt 840  
aaacaaagca gttctctcct tcgtgtaagc agaaattgag atctctttcc agtgacctca 900  
gaggggaagg cggaggagat tgggtgctgg ccaacctgtg tccacttgcc tggctgagca 960  
cagacatttg tagggcctcc ctggacagga acaagcatcc tgccagcagt ttctgatgag 1020  
gaateccttc attttccagc agcttgacga ggaggccagg ggaaaatatg gtgaccacag 1080  
aaggcctgga tcatgagggt tgggagtaaa gctctcagat gaagttgtca aagatggcac 1140  
catggccttc gttgttgagg atgggatatg cctcagtggt ttactagatc ctgattgtta 1200  
tttgggaact ccgggatgtg atcattcctg aggtttcttc cagctctcaa atcccaagaa 1260  
agcatttggt acagattaaa gcaggacagc aaaaacatgc cgcaccctgg tcctacatca 1320  
ctaaaataaa tcccaaagta ctgagctggg atctggcctc agaaggcttc tcaaagccct 1380  
gtgctgtgcc gctgccatga attgtagctg atttatgaga taaaacctaa ttaccaa 1438

<210> 58  
<211> 901  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:980769.1:2000SEP08

<400> 58  
ccgcgttcat aacttcaatg gcattttggc ttgtgagcac caggttacca gaatacaaga 60  
agtccaagat gactgaaaaa ccttgaactg cagcaatatc taagtgggta gtattgtttt 120  
ggttagggct ttctttgttt gaaaagcaat ataaagtctt aaagaaacgg ctgcctgcaa 180  
ccaggatggt cttatgagct ttgaagattt ttccgcttac cacaatgctg acatcacaaa 240  
gaataccttt ctttctctgt tcatttagtt gtgcgagaag ttgatagcaa taagagttct 300  
catttggcct actattaaag tcacagtacc cctcctctga tggatatttct gactcctggg 360  
cgttggcttc gtggagtttg tgctcggtag agacgtgggt cctcaggagg agggttttct 420  
tgtagttgcg ggtccctcgc gaaagctcat catgccccac tgggggaagg tccacgcttc 480  
ccccatctgc aggggtcttt tggcaccagc cgcaagacat aagaccagta tccttggaa 540  
agcgcagcca gggaaagtcc ttgagccagg aggtctggaa ggagcattgg agcttctgca 600  
ttagacactt gggccgcgcg agcgggaagt gctggacagt ctctccttcg ccagctccca 660  
caccgctgcc ctccacttcg tccccgccat cgccgcccga ccgcctaagt gctgctgcac 720  
gtgcgcctt ccgccccgct gccgtcgccc aggtctgccc cctcctcctc gtccctgggt 780  
ctgtcgctca gggacgcgcc gtccctgcat aactccttcc cctccagcgc gtccagctcc 840  
agatccacag tcgccggccc tcgccattg aaatgcctca agggccacac cgaggtctcg 900  
g 901

<210> 59



&lt;211&gt; 1088

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:332474.3:2000SEP08

&lt;400&gt; 59

```
ccggacccca gactcttcct acggcgactg acacctcaga caccgcttgt ggggggaggc 60
gctctggctc cggggctgtg cccagcgcgga gcgtggccgg caccacttgc ctgcagagcc 120
ggcactgcac tggggctcctt gggggcggtt cctccagacg tctctgaggg ttgtcgggtg 180
ctgccagatt ggggtggggg acgctacagg tagtgggtcc ggagcagcgg ctggcccttc 240
atctcactgc ccttctcggg accccaccgt cctctggaag cttgtgcagg gcacctgcca 300
ctgtgaccac ctggacgcag tgggttgttt tttccgtttc catctctggc tagtcttgaa 360
ggctagactt gcgtaggaga gcagacacga ggacatgagc cagttctgct gagctgttac 420
tgtatacctg ctttcccaca acagcaagga agaaccatgg cccaggggtc actgtcattc 480
ggggacgtgg ctgtgggctt caccgggaaa gagtggcagc agctggacct ggagcagagg 540
accctgtacc aggatgtgat gctggagaac tacagccacc tgctctctgt gggctgaact 600
gtctgcttgc ttggaaggcc cagctggctc agcacaatc ctttaaccgt ttctcccaa 660
gaggggtgtc agtcagcaaa ccagctgtga tctccagttt ggagcagggg aaggagccat 720
ggatggagga ggaagagata aggaagtgga gcttcccaga aagctgactt catgagaccg 780
aggatctcgt tcgttgatcg atgctacata ttcagcatgt tagataatga ggttggtgag 840
attgatgaaa tcccgccac cagggccgac ccacaaagag ccactttatg aagaacatgc 900
tcctgaactc cactccacaa ccagcttttg aaaagaaaaa tctgaaaaaa tttgacctca 960
gttaaactgg tccaaatgcc tgagcataac caaaatcttt cctcattcac tcttcacgtg 1020
gaccgtgttt aatttggttg ctggtgatca tatttgtttt agtaaaggct gtcactctat 1080
cttgggagc 1088
```

&lt;210&gt; 60

&lt;211&gt; 1231

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1087707.1:2000SEP08

&lt;400&gt; 60

```
tttttataat cagaaagtga tataatctta tatacaacgt taaatacaaa tgtccctcag 60
gcagggccag gcccacaaca caaggagaaa tcgctgcctg aacatgctgc aggcaaaaac 120
ctgtcactct ttctcattc agcccagtg ttcattacat ctctgtcga tcagggtcta 180
aatgtgtggg gcggagaacc ccagccaatc agtgggtgcta gcgtgaaaac tgcccaatca 240
ggtgcccagc tagagaggaa gaggggggct cttcaatatg gcaaggcctt tgtctcctag 300
cttctaggct ctgagtcag taccgctctg tactattcca tctcttcgc tccattagct 360
cctcggtgac tccaccatag cccctgttat cctgtgacct gcaggctact ggagatccat 420
agggagaaga gcggaacatc cggaggctgg gaaatgcgag tgctaacatt cagggatgtg 480
gctgtagaat tctcccaga agagtgggaa tgcctggact ctgctcagca gcgtttgtat 540
agggatgtga tgtagagaa ctacggaaac ctgttctcct tgggtcttgc tatctttaag 600
ccagacttga ttacctatct ggagcaaaaga aaagagcctt ggaatgcaag gagacagaag 660
acagtagcca aacaccaggg ttattatgat gtttgcctg aagattatga gtataattgg 720
agctatatgt ttctgaattc tgaacaacta tttacaaaat tttatcctac tttttctgt 780
tgaacatatg acttctctgg tctgctaaac acatacagac ctttagtttt ggtttacatg 840
gatttaaata tatagatata tcaactgtaaa ataaacttca ggtgtaacag atttatagag 900
aaagtaatac taattgttta tggttgtgta cctactttga gaagaaaaga aaaatattag 960
aatgaacaga taattttaca agtgttgatc acttaccagc caaccagaa acttcagaga 1020
ttttgaaagc aaatctatct tctctgctgt gtattaaatt catttatcta aaatgttatt 1080
gctcctggct tagaatcatc ttgtgcaaat tctctttttt tggttgttgt ctgtttgcct 1140
gttgctcacc atagacataa ttttcttttc ataaaacatt ctttgataaa tcacctcaga 1200
gattatgaaa gtgactttga taaaatttaa t 1231
```

&lt;210&gt; 61

&lt;211&gt; 1024

<212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:415349.1:2000SEP08

<400> 61  
 cgagtgggtg agaaagtgggt gtctgattta acaaaataat attgggagtc agatgtgttg 60  
 gagctgaaga taccaacaat ttttcacttc cttgtgtttg gctcaaagtt gcctgattca 120  
 caggagtagg aagtttctga gtaggtgtga tatttggtta agggggcgag ttattgctag 180  
 cactaggtgc aacatctgaa atagcaacag ggcctgggtt agattggacc tgtgccactg 240  
 tgttattact cggcaaagtc tcctttgttg gcagaatctc ggagcaaaaa atgagattct 300  
 gatgatgatg atgatgatga tgggtaacaa tgatcttttt catttcataa tcttcggcag 360  
 ataataaaaa agactctgtt ataataaggca ttatagtagc ccattatctt tgggcttcat 420  
 ctttttgattt ctgtattaca aggttcttat caccagaatc aggaggtaaa ggctcagtat 480  
 taccatcctg cgctgtacct gagatgcttt taacctgtga caatgggaca ccaagctctg 540  
 ctataaattt cactcctaata aactgccctg atttaattaa ctcatcaagc aaatctgac 600  
 taacacgaac aatttttagaa ctatagatat aattgagaat ttctgcaaag atctctgctc 660  
 ttataaagct cagttcaaca acttgcccag caacagagaa gagctgatgg aagtaggtac 720  
 tagaagctga aagaatattc ttgtgagccc ggaattttcg gtcttccaca ataacggtaa 780  
 catcacagaa gagtccatgg ccacgttgct cattcaagga gtccagcaga ctgccagagt 840  
 actgaatgtc tgtagcagaa atcagttttg tactctccat gcctgcttcc cggcgccgcc 900  
 gcttctcctc ctccggacgc gcggaaggct gggagtcccc caacaagaag aagccgcact 960  
 cccacgcat tctgcccctt ctttaaacgt cgctccagta tgggtcgacg tcgacgcgcg 1020  
 gccg 1024

<210> 62  
 <211> 1540  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:132420.2:2000SEP08

<400> 62  
 cagggtgggtg tcagtgtcac atacaagagt aaaactgttc ctagagcaaa tggcgatgag 60  
 agagctgaac gcggattcat gctcaagccc ccagatgggg gctatgtggg agactagtgg 120  
 gtctgtgaaa gagaactcca gtcagagtaa gaaatacagc acaaaaatag agaactctgg 180  
 tcctgagagc gcttgacggc acttctggag cttccgttat catgaagcaa ccggaccgct 240  
 tgagactatc agccaacttc agaaattgtg ccatacagtg ctgaggccag agatccactc 300  
 aaaagagcag atcttggaat tgctgggtgt agagcagttc ctgagcattc tgcccaagga 360  
 gacccagaac tgggtgcaga agcatcatcc acagaatgtc aaacagggtc tggctcctgt 420  
 ggaattcttg cagagggagc ctgatggaac aaagaatgag gtaagaagaa agattgccta 480  
 catgtacaat gaagtaggat aaaagtggat atatcagaat ggaggagtta tcagagaatt 540  
 agacagatta agtcgctctt ctgttgctcc tttaggacag ctagcaagga tctgaaatga 600  
 catacatgca tcaatagtcc acaggtttca agatttctgg cagagccttt ccttgtctta 660  
 gagaaaataa atgcattttg tcttctctag ggccagggtg aattgatggg ggtgggccta 720  
 gtctggcatg ggcaaggaca agagttgagt tcttagaaaa agagaaataa agggcaagaa 780  
 agagttggga aaggggagaa gatccacaag atgggcagca cttacagaga ccaaaggaga 840  
 gagaggttag caaacagaag tatgaatcaa gcggctcttg ggagactgca aagaaagaga 900  
 acaaatgga acagaaaaaa gggaaacaag gatattgaaa gacgacaatt gccacagag 960  
 aagtggaggg aatcaggatg ctaaggggta gtatcttcta tagcagaaac acacatctct 1020  
 gggctgggag gaaagagcct ccattgccct cacttagtgg cttgtcattg gcccttgctg 1080  
 gtgagtgga aaccaagggt tgctccttga agatggcagc acactttctt ttttcaggag 1140  
 cctaccttga aaaatccagg gcaggccata tttttggtgt cttattacta ccaaatcatg 1200  
 tgaagagaga ttagagccca actcttctcc ccatttatc ttccatccat ctccattccc 1260  
 tgtgaaattc tttcctggag gattggaaac tggctaacag gaaagtttat ttcactcaat 1320  
 tccttttctt ttacagaaaa ccaaaagata aaaattcact aaaacaagga tccaagtcaa 1380  
 gaaagtaaac cagtgaaggt tcccagcag gggacaggc tgcctacca actctgctag 1440  
 tggcattgagc cagggctctg tgcatctggg tgggggccac tgggttgggg ggcggggatc 1500  
 tgcagtagtc acagcttttt ctattttatc ctaattatcc 1540

<210> 63  
<211> 642  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:394201.1:2000SEP08

<400> 63  
atgactgctg aaattgcaat agtgattgca gagttccctg catcactgta ttggtggcat 60  
cttttactag ttgtgctttt ttgatgttcc acaggaatga gcatagatga acaatgttgc 120  
atcacgcttg acatcacggg ggcttctttc cagtgtggat ttgctcatgt aaaatgaggt 180  
gtgagctctg cctgaaagct tttccatatg catcacattt gcagggcttt tctccagtgt 240  
gggttctttg gtgtctcaaa agatgtgagc tgttactgaa agctttccca cacacatcac 300  
actcataggg cttctctcta ccgtggattc gctggtgtcc aacaagagct gaactgtatc 360  
tgaaggcctt tccacgcttg tcacattcat atagtttctt tccactgtgg attttctggg 420  
gacagaagag gcccaagcac tagctaaagc ttttccctca ctactacac tgctatgggt 480  
tctcttcagt atgaactctg atgttgtctc agatatgaac tcagagagga tttcccacaa 540  
tcattacact ggtatgggtt cttttccgtg tgagttctct ggtgtcaaat aacatctgag 600  
ctgtgatgaa agactttccc acactcacta cattggaaac gg 642

<210> 64  
<211> 1019  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1060884.1:2000SEP08

<400> 64  
ggatttttaa tatctgcta atacttatgc cttcatcgca ttaatgtgct acgcttgatt 60  
taattgttcc cttttgatat tcagtttatt tccagtttga tgctgcaata ccactgtgct 120  
gaacgtctct ggcacattca tcttggttgc aattgtactt taggctatat ttctaagagt 180  
gaaatgacca cagccaaagg tatgaacatt gttcaggctt tcaataaaga ctgccagata 240  
tttacaaga ttgtgccaat ctctattctc accaccgctt ttgattattg tcttatctc 300  
tgcttgccac tgagccccgg gccatgtctc tgtaactcac taacggccac atccccattt 360  
cttccccatg agtaggattt cagtttccca aacctgagat gatctgtcag ctggagaact 420  
gggacgagca gtggatcctg gatctaccga gaactgggaa taggaaggct tccggtagtg 480  
cttgcccagg ttctgaagcc agacacaaga tgaaaaagct aactccaaaa cagaaatttt 540  
ctgaagattt agagtcatat aagatatcag tggtaatgca ggaatcagct gagaaacttt 600  
cagaaaagtt acataagtgt aaagaatttg tggacagttg caggcttact ttccctacta 660  
gtggtgatga atacagcagg ggcttccctc aaaaccttaa cttatttcaa gatcagaatg 720  
cgcaaaacaag gtggaagcag ggcagatatg atgaggatgg caaaccttc aatcaaagat 780  
ctttgtcttt ggggcatgag cgaattctca caagagcaaa gtcttatgaa tgcagtgaat 840  
gtggaaaagt cattagggct aaggcatggt ttgatcaaca tcaaagaatt cactttttag 900  
agaatccttt tgagtgtgag gtctgtgggc aagccttcag acagcgggtc gctcttacgg 960  
tcccataaac agtgtcacct gcaaaacaag ccatacagat gtcatgactg tgggaagtg 1019

<210> 65  
<211> 1658  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:242191.1:2000SEP08

<220>  
<221> unsure  
<222> 29, 58  
<223> a, t, c, g, or other

&lt;400&gt; 65

```

ctcagcagct ttctcaccac acacctganc tgcgggccagt ttgcagtacc aaggaganct 60
gctgggaccag agatggagac tgaggtccca gaaaggcact gctggagcag gaagatctgg 120
gccagccacc aaagggatgt gaacctggag cgaagcagga ggactgaaac ttgggcaacc 180
ttcctcttca ggggtgctgg agggctcctg ggcacatcca ggccggggaa gaagcccccg 240
ggagccgtgt ggtgttcaca cgcatacagc gttcaggaga atgcaggagg gggcctgaga 300
agagctgcgg atttgggtcac tcccgggaag cgttaggtag agtggagagc agagccagac 360
cgcgggggctc caggaagtgg agatcctgtg gttaaataga cgttgccgca cgtggagctc 420
aatagacaat tcttgttttg actttocagg agaggagtgg atgattcgga aggtgaaggt 480
ggaggacgaa gatcaggagg cagaagagga ggtcgaatgg ccccagcatc tatcgttact 540
tcccagcccc ttcccgcgcc ctgacctggg gcatctggct gccgcgtaca aactggagcc 600
agggggccccg ggggcaactga gtggggtcgc gctgtctggg tgggggtccga tgccggagaa 660
gcctacggcg tgcggggagtg gtgagcggcg cttccgggac cagctgacgt tgcgactgca 720
ccagcggctg caccggggcg agggcccctg cgctgcccc gactgcgggc gcagcttcac 780
gcagcgcgcc cacatgctac tgcatacagc cagccaccgc ggcgagcggc ctttcccgtg 840
ctccgagtgcc gacaagcgct tcagcaagaa ggcccacttg acccgccacc tgcgcacgca 900
cacgggagag cgcccttacc cgtgcgcgga gtgcggcaag cgcttcagcc agaagatcca 960
cctgggctcg caccaaaaga cccacaccgg cgagcggccc ttcccctgca cggaatgcca 1020
gaagcgcttt cgcaagaaga cgcacttgat tcggcaccag cgcattccata cgggagagag 1080
gcctaccag tgcgcacagt gcgcacgcag cttcacgcac aagcagcact tgggtgcggca 1140
ccaaagggtg caccagacgg ccggcccggc caggcccctt cccgactcgt ccgcttctcc 1200
tcattccact gcccgctccc cgaccccatc ctttcccggg ccaaagcctt tcgcgtgctc 1260
cgactgcgcc ttgagcttgc gctggaaaaa gaacctcgcc acgcaccagt gtctgcaccg 1320
cagcgagggt cgcccccttg ggtgcgatga gtgcgcactg ggcgccaacc tggatgccc 1380
cgccgccaa cccctggcca gcgcgcctgg cggaccgggc tgcggcccag gatccgatac 1440
cgtggtgccc cagcgcgccc cctcgggcca gcggtccttc ttctgcccgg actcggggcg 1500
cggcttctcc catgggcagc acctggcgcg gcaccgcgc gtgcacacgg gcgaacggcc 1560
cttcgcctgc acgcagtgtg accgcccgtt cggtctcgcg cctaactctg tcgcccactc 1620
cagggcccac agcggcgcca ggcctttcgc ctgcgctc 1658

```

&lt;210&gt; 66

&lt;211&gt; 1350

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1063762.3:2000SEP08

&lt;400&gt; 66

```

caggcgcaact tccagttctg cccacccgct gcggccattg tccgaccccg gtgcgggtga 60
ggcccccttg ggcagcccct ccgcagatca gaattggaga caaccgagcc ttcggcgggg 120
gcgggaggag ctgcccagag ctctgggtgg gccggaggct gcgaaatccg gagcccccca 180
gaggcggtga ttctgactcc tgccccggag ccggggccctg gcgaggcagg aatggccccg 240
aggcctccga ccgcgcgccc ccaggaatca gtgacattca aagatgtgtc tgtggacttc 300
accaggaag aatgggtacca tgtcgacctt gctcagagga gcttatacag ggatgtgatg 360
ctggagaact atagccacct ggtttctctt ggatatcaag tttccaagcc agaggtgatc 420
ttcaaattgg agcaaggaga agagccatgg atatcagagg gagaaatcca acgacctttc 480
tatccagact ggaagaccag gcctgaagtc aaatcatcac atttgagca ggatgtatca 540
gaagtatccc actgcacaca tgatctctta catgctacat tagaagactc ctgggatggt 600
agcagccagt tagacaggca acaggaaaac tgggaagagac atctgggatc agaggcatcc 660
accagaaga aaataattac accacaagaa aattttgagc aaaataaatt tgggtgaaat 720
tctagattga acaccaatth gggtacacaa ctgaacattc ctgcaagaat aaggcctagt 780
gaatgtgaga cctttggaag caatthggga cataatgcag acttacttaa tgagaataat 840
attcttgcaa aaaagaaacc ctataagtgt gataaatgta gaaaagcctt tattcataga 900
tcacgctta ctaaacatga gaaaacacat aaaggagagg gagctttccc taatggaaca 960
gatcaaggaa tttatcctgg aaagaaacac catgaatgta ccgactgtgg gaaaaccttt 1020
ctctggaaga cacagcttac tgagcatcag agaattcaca ctggggagaa accctttgaa 1080
tgcaatgtat gtggaaaggc cttcaggcat agctcatctc ttgggtcagca tgagaatgct 1140
cataccggag agaaacccta tcagtgtagt ctctgtggga aagccttcca gcgcagttcc 1200
tcctttgttc aacaccagcg aattoacact ggagagaaac cctatcgatg taaatctatg 1260
tgggagggtcc tttaggcatg gcacatccct cactcaacac gagggtcaca cacagtggag 1320
agaagccctt ccagtgtgaa gatgtgggaa 1350

```

<210> 67  
<211> 545  
<212> DNA  
<213> Homo sapiens  
  
<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1100856.1:2000SEP08

<400> 67  
ctctctctct cttatcatgg gtcgctcccg gcggaccggc gcgcaccgag cacattcctt 60  
agccccccag atgaaggcca agaagcggcg gccggacctg gatgagatcc atccggagct 120  
gagcccgag gctctcccg gtccaaagcg ggagcgggac gcagagccgg atccggacct 180  
gccagggggc ggcctgcac gctgtctggc ctgcgcgagg tacttcatcg attctgcca 240  
cctgaagacc cacttccgat ccaaagacca caagaaaagg ctgaagcagc tgagtgtcga 300  
accctacagt caggaagagg ctgagagggc cgcgggcatg ggctcatatg tgcaacctca 360  
gcgactgggc gtgcccacag aagtgtccac tgagatccct gagatggaca cgtccacctg 420  
agtggccaaa ggattcagga cagagaaact gcccatgggc agtgtgggtc attggcaggc 480  
tgagagtgtg tatgtgggtt tgggggtgca gacatgccct cttctgggtg ttctgccaaa 540  
ataaaa 545

<210> 68  
<211> 1399  
<212> DNA  
<213> Homo sapiens  
  
<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:979390.2:2000SEP08

<220>  
<221> unsure  
<222> 755  
<223> a, t, c, g, or other

<400> 68  
aattcactgt ctgtagcatc tgctcctcca cagagggacc ctggaatggc gatggcactc 60  
ccgatgcctg gacctcagga ggcgggtgtg ttcgaggatg tggctgtgta cttcacaagg 120  
atagagtggg gttgcctggc ccccgaccag caggcactct acagggacgt gatgctggag 180  
aactatggga acctggcctc actaggcttt cttgttgcca aatcagcact gatctcccta 240  
ttggagcaag gagaggagcc gggggccttg attctgcagg tggctgaaca gagcgtggcc 300  
aaagccagcc tgtgcacaga ggacccta atactgcca gcagaagcca gggaagggaag 360  
ccctgccagc ttcagaaggt gggccaggag agaaggggtg ggcctggaag ggtagccggt 420  
gggggtgctg catcctcttg gccccacggg gagcaccctg ttcaccccta acagatgaag 480  
agaaggtgag gggggattga gttacctgcc cactggtgtg tcaaaactcca tgatgggggt 540  
gggaacatgc agacctgggc cctccaaggc ccactgctcc ccgggctgtg cggccacag 600  
tctgtgccag gggctgcaga aagagccatg ggaagcctgg ccctggtagg gcaccaagaa 660  
tggggagggc tggatccacc ctccatagaga acaagaatcc tgggcagtag ccgggacatg 720  
gtgggggagg ctgaagggtt gtccctgagc atccnccaac atggagctcg gaagtctggg 780  
ccatgggcat ggaaagcacg aagggtcagg ccagcctggg gaaccaagag ctgtgccacg 840  
atcccttcca cctgtgggca gctgccaca tgtcccgcga ctgggttcca tcacctgaga 900  
ctcctggcag gtgcctctct gagtctttgt cctgtgctgg aaactgtgat gtgcatgcac 960  
acatcttcat caccacagc acggacagcg ggaataacac ttcccttggg gatccagggt 1020  
ttctggtgtc tccagatttt tcaggcgccc accatgttcc tctcatccaa acgtcggcac 1080  
ccaaggtgga agccaggttc atggcatgcc cggcccagct gcagtcagag tgcagatccc 1140  
atggcgagcg tgggatccag cccaaagcac aagccacgca cagcccgcgc aggttgagtg 1200  
ggcaggatcc tccaacagcc agcccagggt gagtgggagc gtacctcaa cagcctggag 1260  
ccgtgaggcc ttgcgcaggc cggttgcggc cgtggagggt tctggctggc aaagtggcac 1320  
cgaaagagtc ctgtgtcata accagactca cttgtagctt ctgttgctct ccctagctat 1380  
ttgcaaagca tcccaggat 1399

<210> 69  
<211> 629

<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1400447.1:2000SEP08

<400> 69  
ttctcttgag atcgctcagt gaggctgaag ggaaagacca gaggggtgggt ggaaggcatg 60  
agaactggac ctggagcagg ctttctgcag ggaaaagtca atgtatagta ccctcagaag 120  
gaggagagag aatggggagc aaacaggaac caaaacactc actgcttggg gtactggagg 180  
cctcccagga tgtgaatgct cactgcctcg ctctggactg ctgtatgaga gcttttgacc 240  
cagttctctc ctaatagcag gtgtgtggac ccttctagcc tgaggagtcc tgcagggtgtg 300  
aagctccaca cctgcctcca tagcactttg cctgtcccta agagggctca tcggagaaga 360  
aagaatggct gtcagccacc tgccaacat ggtccaggaa tcggtgacct tcaaggatgt 420  
ggctatactg ttcacccagg aagagtgggg gcagctgagc cccgcccaga gggccctgta 480  
cagggacgtg atgctggaga actacagcaa cctgggtctca ctgggactct taggacccaa 540  
accagatacg ttttcccagc tagaaaaaag ggaagtgtgg atgccagagg acacccctgg 600  
aggcttctgt cttgatggag tctcactct 629

<210> 70  
<211> 740  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1400562.1:2000SEP08

<220>  
<221> unsure  
<222> 33  
<223> a, t, c, g, or other

<400> 70  
gtcggtcctt tgtctctcgc tgcagccgga ggncaaggtc tcgtccttac tgcgtgtgtg 60  
cctctgctcc tagagtccca gcctctgtgg ccctatgacc tgcagggtatt gggagatcca 120  
caactaagtc gccaggaccc cctggaagcc tagaaatggg atcattgaca tttagggatg 180  
tggccataga attctctctg gaggagtggc aatgcctgga cactgcacag cagaatttat 240  
atagaaatgt gatgttagag aactacagaa acctgggtctt cctgggtatt gctgccttta 300  
agccagacct gatcattttt ctggaggaag gaaaagagtc ctggaatatg aagagacatg 360  
agatgggtgga agaattccca gttatatgtt ctcatcttgc tcaagatctt tggccagagc 420  
agggcataga agattctttc caaaaagtga tattgagaag atacaagatt catcatcatg 480  
cctgtgagct ggggtccaata atgaatcact atcctacctg tggccaaatg catatatgac 540  
agttacaact ctaactgcag actgcatcgg catgtaaaat ttaggtcacc agtgggctct 600  
ctccatgttt gaatgtgggt attctaattg tcagcgaggt gtgcatagga gagtcacaat 660  
ctcaccttca ggtgggcctt gtatcagcac tctctctact accttaagac tgtatgacac 720  
atgggtgagc taaaaatgtt 740

<210> 71  
<211> 768  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1076130.1:2000SEP08

<220>  
<221> unsure  
<222> 38  
<223> a, t, c, g, or other

&lt;400&gt; 71

```
atcaggtaaa agcaaatgag agtgactgta gatcaganag acagtgggca aagacctcag 60
gggagaaaaag aggaaaactg acactcccgg agaagagctt aagtgaagtc ctaagtcaac 120
agagaccttg cttgggagag agaccctata aatatctcaa atacagcaaa agctttgggtc 180
caaaactccct tctcatgcat cagggtatccc accagggtgga aaatccatat aaatgtgctg 240
attgtgggaa aagcttcagt cggagtgcac gactcattag acaccggaga atccacactg 300
gagagaaacc ttataaatgt cttgactgtg gaaaaagttt ccgtgacagt tcaaatttca 360
tcacccatag gagaatccac acaggagaga aaccttatca atgtggtgag tgtgggaaat 420
gcttcaatca gagctcaagc cttatcattc accagagaac ccacacagga gaaaagccct 480
atcaatgtga agagtgtgga aaaagcttca ataacagttc tcattttagt gcacatcgga 540
ggatacacac aggagagaga ccccatgtgt gtcctgactg tggaaagagt ttcagtaaga 600
gttctgactt acgtgcacat catagaacct acacaggaga gaaaccctat ggggtgtcatg 660
actgtggtaa gtgcttcagt aaaagctctg cccttaataa gcacggagaa atccatgcac 720
gggaaaagct tctgacacag tcagctccca agtaagcccc tgaggata 768
```

&lt;210&gt; 72

&lt;211&gt; 1098

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1064459.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 89, 94

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 72

```
tcagagattg agcaggaacc tgttgtttca ggagtcagtg acctttgagg atgtggctgt 60
ttacttcacc cagaatcaat gggccagcnt cgancctgcg cagagggccc tgtacgggga 120
ggtgatgctg gagaattatg caaatgtggc ttctctggta gcgtttccat tccccaaacc 180
tgctctgac tcccacctgg agagagggga agcaccatgg ggcccagatc cctgggacac 240
cgagattctg agaggcatca gtcaagggtg tgagtcctgg atcaaaaatg aagggctagt 300
tataaagcag gaagcctctg aagaaacaga gttgcacaga atgccagtag gaggacttct 360
caggaacgtt tctcagcact ttgattttta aaggaaggca ctgaagcaga ctttcaatct 420
aaatccaaat ctgatacttc gaggtggaat. gaagtcttat gaatgtaaag aatgtgggaa 480
aatcttccga tataactcaa agcttattcg gcacagatg agtcatactg gggaaaagcc 540
ctttaagtgt aaggagtgtg gcaaagcttt caagtccagc tatgattgta ttgtacatga 600
gaaaaaccac attggagaag gggccctatga atgtaaggag tgtggcaaag gtttgagttc 660
caacacagcc ttgactcaac atcagaggat ccacactgga gagaaaccct atgaatgtaa 720
agagtgtgga aaggcttttc gtaggagtgc ggcatacctg cagcatcaga gattacacac 780
gggagagaaa ctctataaat gtaaggaatg ttggaaagct ttcggttgta ggtcactttt 840
tattgtccat cagagaattc atactgggga gaaaccttat caatgtaagg agtgtggcaa 900
agccttcacc cagaaaatag cctccattca gcacagaga gttcacactg gggagaagcc 960
ttatgaatgt aagggtgtgt ggaaagcctt caaatggtat ggaagttttg ttcagcatca 1020
gaaattgcac cctgtggaga agaagccagt caaggctcct gggccatccc tgggtcagtc 1080
ccagtgtctc tctccagc 1098
```

&lt;210&gt; 73

&lt;211&gt; 370

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1079415.14:2000SEP08

&lt;400&gt; 73

```
aagaccagga atacttcccta gtaagtctta ccatgtttat gccattggat gtttttcccc 60
caggaatgtt ctgcatagca attgtatctt ttgggtttaag ttgagtttcc cagtctgaaa 120
tagcacagga atttttataa ttattgggag aaaaggttgg gatcaaagaa aaagggaggg 180
```

agaccctatg gatttgtttt caaataatat cttctaagta gagaaatgta agaaagaaga 240  
cacaaccagt gattcaccag tgataggaac aagcagtaga aatgggaggt attttagtta 300  
gaaaggattg tggaggttgg gaaataagac agtattatct ggaaaatagg gacaactaaa 360  
caaaaagcat 370

<210> 74

<211> 763

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1329431.3:2000SEP08

<400> 74

gtgcagtggg tgctgtcagt gccagctggc gccgaggaac tcagcggcgt ggggcgagcc 60  
ctggcccttg tgggctcagc gggctcgtgc tgccactgcg gctccagcgt cccctccgta 120  
agccccaagc ctgtggggcc tgggcctggc cgggcggccc agcgtgctc tgggccgagg 180  
gctcctggct cctccgaat cctgtgagg gcgcgcgggg tcttctcag cgggagtcgg 240  
ggtttttagag ctgcggattc cagggtctga aagcagaagg gggtcttcct ggctcccttt 300  
ttcttctcag atctgccttc tggagactgc gccgtcctcc cgggagagcc agaaagagga 360  
catggctgct gggcagcggg aagcagggcc ccaggtgtca ctgacattcg aggatgtggc 420  
tgtgtctctt acctgggatg agtggagaaa gctggctcct tctcagagaa acttgtaccg 480  
ggatgtgatg ctggagaact ataggaacct ggtctcactg ggactctcat ttaccaaacc 540  
aaaagtcac tccctgttgc agcaaggaga agatccctgg gaggtggaga aagacagttc 600  
tggcgtctcc tctctaggat gtaagagcac acctaaaatg acaaagtcaa ctcaaactca 660  
ggattcattt caggagcaga taaggaacag attgccaggg gatgaaccct ggaacttcat 720  
atcagaaaga tcttgcata atgaagagaa attaaagaaa cag 763

<210> 75

<211> 545

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1088431.2:2000SEP08

<400> 75

cgcaggttcc tgccgaccg gaagcggatc tcgcggggct actggcgctc tcggcccaca 60  
caatatgacc tcggggagga tgccaggaag atgaactgtg atgatccact tcttcttaat 120  
gaatgactga cttacctgag aaagaactc agaggaagag gaaagaaaga agaggagga 180  
atggctcttt ctcagggact gtttacattc aaggatgtgg ccatagaatt ctctcaagag 240  
gagtgggagt gcctggacc tgcccagagg gccttgtaca gggacgtgat gttggagaac 300  
tacaggaacc tgctttctct cgatgaggat aacatccctc cagaagatgg ttctcacctt 360  
gcagcctgtg gacagagcac actgcctctt ccttagaatc ctacaaaatc cgaccctttt 420  
attttactcc cgtacctttt atttctctcc cttttctagc tttctactct atccgggaac 480  
ttagggagac cttagtcttc ctgattccat attcttctaa gatttgcctc tgaaagttgc 540  
ttttg 545

<210> 76

<211> 1636

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1329462.2:2000SEP08

<220>

<221> unsure

<222> 556

<223> a, t, c, g, or other



&lt;400&gt; 76

```
gtgcgcgggtg tggctccagg tgaaggatgc cgggaggcct cggcaggacc gcgcggaaac 60
gggcctttctg cccaaaagat gctgctttct tccttattct ttccctcag aatctcgctg 120
tctccttcca accacctgtg gtcggcatcc ccgcgttgct actgcgacgc agaggcgagc 180
gaggtggcgg gaagcaccgc cggggcgggg agggaccctg cgggcgcgga ctccacacca 240
agcctctgct cagcgtcacc ccgcttgctg tgtcctcgca ggtcgcagct tcattggcctg 300
atgccttcag gaagtatttt gaagtcacgc tggctgtgga ttgggggatt tcttgtttcc 360
actgacctgt gaggcgcgc acgtggaggg aggcaccccg ggtcctccgg cactgtccgg 420
cctgcctgt gtccctagta gcagtgggca tttccagacg gtgcagcttg tggctaaagt 480
gacaggaaga ttaggagct ttcagtcttg gatgaggatt cgaactgaag ggcttaggcc 540
cagctgtctt ggagcnaaac atctgttggt gggtgagaag ggaaggagg ggtgagggtg 600
caggggccct cctaggagga catcggaact gttgggcga gacctcatta gagatctaga 660
gcacggatgg gaggagcaga ggctcgattt gtgtcctgtt cagactccca caggctgggt 720
gatggtggga gatttggtgc cattcagatg tggcggcaga ggaggcgcaag gccgaaggag 780
cagacagcac cgcttcttgg ggagtgtgta aggcacatg cggaggggcg agcttagcag 840
ccaagtggag gacagcacc tccatgcctg gattcggtac tcgctcgttc tcgatgttga 900
gctgctggca tattgcagca caactagaga tgtacggatg ccccatctt gatcttacag 960
aatcagaggt gcagccgcaa gaaagctaca cctctcggt tttctctgct ctgccaacac 1020
tagtggaata tgatcacatc ccagggatca gtgtcattta gggatgtgac tatgggcttc 1080
actcaagagg agtggcatca tctggacct gctcagagga cctgttagag gaatgtgatg 1140
ctggagaact acagccacct tgtctcagta gggatttgca ttcctaaacc agaagtgatc 1200
ctcaaattgg agacaggcaa ggagccatgg atattagagg aaaaatttcg aagccagagt 1260
catctgggtg agttagtgc agatgggaatt taaagaatta attaatacca gtgaaacta 1320
ttcaagaatg aagttcaatg agtttaacaa aggtggaaaa tgtttctgtg atgaaagcat 1380
gaaataatc attttgaaga ggaaccttct gaataataa acaatgggaa cagcttctgg 1440
ctgaatgaag acctcatttg gcatcagaag attaaaaatt gggagcaacc ttttgaatac 1500
aatgaatgtg ggaaagcttt cctgagaat tcaactcttc ttgtacataa gagagcttac 1560
acaggacaga aaacatgcaa atatactgaa catgggaaaa cctgttatat gtcatttttt 1620
attactcatc agcaaa 1636
```

&lt;210&gt; 77

&lt;211&gt; 979

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:393468.1:2000SEP08

&lt;400&gt; 77

```
gtatgactct cagatgatcg acctgtgcaa cgtggggttc caattctacc gcagcctgga 60
acactttggg ggcaagcccg tcaagcagga acccattaag cccagcgccg tgtggcccca 120
gccaaagccc actccattcc tgcccacgcc ctacccctac taccctaaag tccacccggg 180
cctcatgttc cccttcttgc tgccctcgct ctgcgcttc cccttcagcc ggcacacctt 240
cctgcccagg cagccccggg aacctctgct gcccggaaa gccgagccc aggagagcga 300
ggagaccaag cagaaggtgg agagggtgga cgtgaacgtg cagatcgatg acagctacta 360
cgtggacgtg ggcggctcgc agaagcgtg gcagttccc acctgcgaga agtcctacac 420
ctccaagtac aacctggtca cccacatcct gggccacagt gggatcaagc cgcacgcgtg 480
cacgcactgc ggaagctct tcaagcagct cagccacctg catacccaca tgctgaccca 540
ccagggcacg cggccccaca agtgccaggt gtgccacaag gccttcaccc agaccagcca 600
cctgaagcgc cacatgatgc agcacagcga ggtgaagccg cacaactgcc gcgtgtgcgg 660
ccgcggttg gcctacccca gcgagctcaa ggcccacgaa gccaaagcag ccagtgggcg 720
cgagaacatc tgtgtggagt gcggcctcga cttccccacc ttggcccagc tgaagagaca 780
cctcaccacg caccggggcc ccatccagta caactgctcc gagtgcgaca agaccttcca 840
gtacccgagc cagctgcaga accacatgat gaagcacaag gacatccggc cctacatctg 900
ctcagagtgt ggcattggat ttgtgcagcc tgacagctca agcagcactc cctcgaccac 960
aagggtgtga aggagcata 979
```

&lt;210&gt; 78

&lt;211&gt; 591

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:722577.1:2000SEP08

&lt;400&gt; 78

```
gtgcggaggc gcggtctgctc ctgctcgtcg cgacgtcgaa tttcagcgcc gcccgcgcca 60
tccttggtct cttcgcccggt gaggagaggg gcgacaatgg ctgggaaagg cggggccgacc 120
aacctggaga aggagcagat gtttgggatg gcggagaagg agatggagta cagggttgat 180
cttttcaaca ggcttacaaa gacctgtttt gagaagtgcg ttgaaaaaag gtataaagaa 240
gctgagctca atatgggtga gaacagttgc attgatcgat gtgtttcaaa gtattggcag 300
gtgactaatt tagtaggtca gatgcttggg aaccagcctc agatgtgaac cgagatgcac 360
tagttccatt ttgtagcaga atcctattgc tctgtttatt ctttttgggt ggtatttgaa 420
tcccggcacc tgctattaag tacaccaata cagcaccaga cttgttaatg cgtggatcgc 480
tctgaagaaa aatcacgggtg tcttgccctt tccagtgatg tgcattatct ttgaatcatg 540
aacggggaac acgcctgtaa taaattttat agagctcgag agtgatcctg t 591
```

&lt;210&gt; 79

&lt;211&gt; 703

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:322783.16:2000SEP08

&lt;400&gt; 79

```
ccagagagaa aagactgcga ggtggccgca gctgtggccg gagagcacia agaataaacc 60
agcagtggaa gagaaaatac tgtaagctgg ctgactgctg gtgaagaaaa tgctttattt 120
ttgtggcagg catctgtggg atctgtaata gaaatatatt ggagtaattc aagattctgt 180
ggttggccct tttgactgct ctctctacag gtttaatttg ggcatttact cattttcatg 240
gctccaagga ccatgtatgt gttggggatc ttcaatattc atgttatttt ctcccttgggt 300
cttatatgat tgttaccttt atgaagcttt agtgattaca aagcactttt tttgtccatt 360
tttacctgag ctttgtaaac tctgatttgc aggatggctg gctgtgggtga aattgatcat 420
tcaataaaca tgcttctctc aaacaggaaa gcgaacgagt cctgttctaa tactgcacct 480
tctttaacgc tccctgaatg tgccatttgt ctgcaaacat gtgttcatcc agtcagtctg 540
ccctgtaagc acgttttctg ctatctatgt gtaaaaggag cttcatggct tggaaagcgg 600
tgtgtctttt gtcgacaaga aattcccagag gatttccctg acaagccaac cttgttgtca 660
ccagaagaac tcaaggcagc aagtagagga catggtgaat atg 703
```

&lt;210&gt; 80

&lt;211&gt; 571

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:901355.2:2000SEP08

&lt;400&gt; 80

```
cccagtcctt ccatctggga ggccaaggcg gcttcgcgtt ctgagaatag acagaacctc 60
tgttactctg tgaccggcag gcaccgggag atccgtagct cagacgccag gacatcccgg 120
aagctgggaa atggtgaatg tgccaggctc ggggtcccca gaagaggag agggggcggtc 180
ggaagcggcg ggaaccctct ttgaggctac tgaacatctg actggctggg cacctgctcc 240
tgtcccttca gttgaggggac tggtgacatt cagggatgtg gccatagaat tctctcgga 300
ggagtgggaa caccctggact cagatcagaa gcttttatat ggggatgtga tgttagagaa 360
ctacggaaac ctgggtctctc tgggtctcgc tgtctctaag ccggacctga tcacctttt 420
ggagcaaagg aaagagccct ggaatgtgaa gagtgcagag acagtagcca tccagccaga 480
tatcttttct catgatactc aaggcctctt aagaaagaag cttatagaag catcattcca 540
aaaagtgata ttggatggat atggggagct g 571
```

&lt;210&gt; 81

&lt;211&gt; 907

&lt;212&gt; DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:038859.2:2000SEP08

<400> 81

```
ccgcgttcat aacttcaatg gcattttggc ttgtgagcac caggttacca gaatacaaga 60
agtccaagat gactgaaaaa ccttgaactg cagcaatatc taagtgggta gtattgtttt 120
ggttagggct ttctttgttt gaaaagcaat ataaagtctt aaagaaacgg ctgcctgcaa 180
ccaggatggt cttatgagct ttgaagattt ttccgcttac cacaatgctg acatcacaaa 240
gaataccttt ctttctctgt tcathtagtt gtgcgagaag ttgatagcaa taagagttct 300
catttggcct actattaaag tcacagtacc cctcctctga tggattttct gactcctggg 360
cgttggcttc gtggagtttg tgctcggtag agacgtggtg cctcaggagg agggttttct 420
tgtagtggcg ggtccctcgc gaaagctcat catgccccac tgggggaagg tccacgcttc 480
ccccatctgc aggggtcttt tggcaccagc cgcaagacat aagaccagta tccttggaa 540
agcgcagcca gggaaagtcc ttgagccagg aggtctggaa ggagcattgg agcttctgca 600
ttagagactt gggccgcgcg agcgggaagt gctggacagt ctctccttcg ccagctccca 660
cacgcctgcc ctccacttcg tccccgccat cgccgcccga ccgcctgctg ctgctgcagt 720
gcgcgcttcg gccccgtcgc cgtcgcccag gctcgcccc tcctcctcgt cctcggtgct 780
gtcgtcagg gacgcgccgt cctgcatcaa ctccctcccc tccagcgctg ccagctccag 840
atccacagtc gccggccctc gccattgaa atgcctcaag ggccacaccg aggtctcggg 900
gcggccg 907
```

<210> 82

<211> 543

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1046117.1:2000SEP08

<400> 82

```
ctctgcctca cctgcctcct tgaaggaaac acggggaagc caggagtggc cgtcacgttg 60
gtgacaaata tgtcccagga ctcggtgacc ttgcgagacg tggctgtgaa cttcaccaaa 120
gaggagtggg cctgctgga cccagctcag aggaatctct acagagacgt gatgctggaa 180
aattctagga acttggcatt catagattgg gcaactccat gtaaaaccaa agacgcaacc 240
cctcagccgg atattcttcc taaaagaaca ttctctgaag ccaacagagt gtgtctcacg 300
agcatcagat toccagcact ccacattaag agaagactgg agatgcccc aaacagagga 360
accacacagg cagggggtga ataattgtga gccacctgca gattgcccc tgaggaaaga 420
tgaatctcct gttagcattt gtgacgatca tgaaatgagg aaccacgtct aaacctacct 480
gcaggcttgt gccttcacag ggagattcca taagacaatg taccctaaca cgtgactcaa 540
gta 543
```

<210> 83

<211> 256

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:801015.1:2000SEP08

<400> 83

```
ggaagccgga aaatggactc agtggccttt gaggatgtgg cagtgaactt caccagagag 60
gagtgggctt tgctggatcc ttggcagaaa aaactctaca gagatgtgat gctggaaacc 120
tataggaacc tggcttcagt aggtgatgac gacaacattc cttcacttag agaacaagtt 180
gcccatcaac gatatttcaa gacctggcat gtggaaaggg aatacttcag taaataaacc 240
aagcatggtg acagct 256
```

<210> 84

<211> 715

<212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1175590.1:2000SEP08

<220>  
 <221> unsure  
 <222> 288  
 <223> a, t, c, g, or other

<400> 84  
 tcagttgtcc caaatccttt cgcgacatg tgatgatgca cgccggaggg agaccgtatg 60  
 agtgcaagca ctgtgggaaa gccttcaggt gtcagaaatc ctttcgagtc catatgatca 120  
 tgcacgccgg agggagaccg tatgagtgca agcagtgagg gaaagcctac tgctgggcaa 180  
 catcctttca acgacacgtg agaattcaca acggggagaa accctataaa tgtggaaaat 240  
 gcgggaaaagc attcggttgg ccctcatcct tacacaaaaca cgcgagancg catgctagaa 300  
 agaaaacctgt gagtgggggc agcgtgggaa agtcttcccc cgaggccctc gccctccac 360  
 agatgtcaaa tcacaaacta gagagaaagt ctataaatgt gaaacgtgtg ggaaaacgta 420  
 tggtttggtcc tcatctttac acaaactga gagaaagcac actggggaga aacctgtaaa 480  
 tgcagccagt gtgggaaaac cttcagggcg gctttgtctt tccaaaaatg taagaacgca 540  
 gattggacag aagcccagc aaatgcgaaa aatgtgggaa agctttcagt tgtcccaaag 600  
 cctttcaagg tcatgtgaga agtcacacag gaaagaaatc ctgtacatct aagtaatggg 660  
 ggaaaacctg tgcaaattaa ttcacatata tggttcagaa aattcacagc aggag 715

<210> 85  
 <211> 788  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1170585.2:2000SEP08

<400> 85  
 gtcttctttc tggcattggc gtttttttgg tcttacgccc taggggtggg agaggactcc 60  
 ggtgcctccg ggtctgggct gggccgcaaa gggcggggccc tgggcccggc gcgggcccggg 120  
 agcagaggcc gagaacctgc aggtgattcc tagaggctgc cagctaaggt cccttatgat 180  
 ttctgtctctg gctttgcagt ttccagcctt tcccaagagc agcaggaaat gaacaagtcc 240  
 caggggacccg tgacattaaa ggatgttatt gtggaattca ccaaggaaga atggaagtta 300  
 ctgacccctg ctcagaggac tctgtataag gatgtgatgc tggaaaaacta tagtcacctt 360  
 gtctcagtgg gttaccatgt gaataagcca aatgcagctc tcaagttgaa gcaaggaaaa 420  
 gagccatgga tattagaagt agaatttcca catcggggct tcccgcgaag acctatggag 480  
 cattcatgat ctagaagcaa gataccagga aagccaagct ggaaattcaa ggaatggaga 540  
 actcacaaaa catcagaaaa ctcataccac agagaaagcc tgtgaatgta aggaatgtgg 600  
 gaagttcttc tgccagaagt ctgccctcat agtacatcag catactcact caaagggcaa 660  
 atcctatgac tgtgataaat gtgggaaatc tttctctaaa aatgaagacc tcataagaca 720  
 tcagaaaatt cacacgagag ataaaaccta tgagtgtaaa gaatgtaaga aaatatatta 780  
 ccaccaaa 788

<210> 86  
 <211> 207  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:719531.2:2000SEP08

<400> 86  
 cagcaactgt ttatcatgaa tacaggatgt aggcaaactc aactgcgcc tgccaccaaa 60  
 aagtttgctc agggccatca ctccctgggt ctggggtcct tgaagttata tactgggaaa 120

tctagcacct attgttcgaa ggatgcagtc tcacaagcct gctgtgaacc caacggctga 180  
gtgacaatta cctgacaatc acccccc 207

<210> 87  
<211> 245  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:794623.1:2000SEP08

<400> 87  
agtggatccc gtggaatgac ggtcacgccg cgggcgggcg attgacttct aaagactgtt 60  
ggtacgtgag aaagaaaccc agaagaggaa gaggaagca aaggagtcag ggatggctct 120  
tcctcagggc ctattgacat tcagggacgt ggccatagaa ttctctcagg aggagtggaa 180  
atgcctggac cctgctcaga ggactttata cagggacgtg atgctggaga attacaggaa 240  
cctcg 245

<210> 88  
<211> 778  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1173119.1:2000SEP08

<400> 88  
ccgagcaggg actgtacacg tgtccagcac atcttcacca gcacaaaag gagcagatta 60  
gagagaaact ttctagaggg gatggaggaa gaccgacatt tgtgaagaac cacagagttc 120  
acatggcagg gaagaccttc ttgtgcagtg aatgtgggaa agccttttagc cacaacata 180  
aactttctga ccatcagaaa atccacactg gagaaagaac ttataagtgc agcaaattgtg 240  
ggatattgtt tatggaaaagg tccacactca atagacatca gagaactcac actggagaaa 300  
ggccttatga gtgcaatgaa tgtgggaaag cctttctttg taagtctcac cttgttcgtc 360  
accagacaat ccactctgga gaaaggcctt atgagtgcag tgaatgtggg aaattgttta 420  
tgtggagttc cacactcatt acacatcaga gggttcacac tggaaagagg ccttatgggt 480  
gcagtgaatg tgggaagtgc tttaagtgc actcaaact ctttaggcat tacagaattc 540  
atacaggaag aaggtcttat ggttgcagtg aatgtgggaa attctttatg gaaaggtcta 600  
acatcagtag acatcagaga gttcacactg gagaaaggcc ttatgagtgc aatgaattgtg 660  
ggaaattctt cagcttgaaa tccgtcctca ttcaacacca aagagttcac actggagaaa 720  
ggccttatga tgcagtgagt gtggcaaatc ctttcgccag gtattattga gtgttttc 778

<210> 89  
<211> 778  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1093285.1:2000SEP08

<400> 89  
acatcagaaa attcatactg gaagtagaga actttacaaa tgtaatgaat gtgtaaagg 60  
cgttattcga aattcgaca tagcacaaca ttggagaatt atacacgaga gagagaacat 120  
tacaattgta atgaatgtgg aaaagtattt aatgaacttt caaatcttgc aagacataga 180  
agaattcata ctggagagaa gctctttaca taatgtaatg aatgatggat aaagcattta 240  
gtagagatat tctaggcctt tactagcgcg catctttgta atcgcgacaa tggagagaag 300  
ccttacaaat gtagtgaatg tggcaaggca ttcagacaca agttatcact aaccaatcat 360  
cagagaatcc atactggaga aagaccttac aaatgtaatg aatgtggcaa ggtcttcaat 420  
cgaattgcac accttgacg acatcggaat attcatactg gagagaaacc ttacaaatgt 480  
aatgagtgtg gcaaggcctt tagtcgcatt tcatacctag cacaacattg gacaattcat 540  
atgggataga aactacaaat gcaacaaatg cgtcaaagaa tttagtgtgc actcaagcct 600

tactacccat cttttattcc atactgcaaa gaaatthttgc aaatgtaaag aatatgacaa 660  
ggtcttcaaa cacaagtttt tctaataact cattagagaa tttatactgg agagacttca 720  
caattataat aaatgtgtgg aaaagtcttc aaaaaaattt cacaccttgc aaaaggag 778

<210> 90  
<211> 476  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1091881.1:2000SEP08

<400> 90  
gctgggttcc cgtctgctgc ctctcggaga gtcccggtg actgccgcag gctccatcgc 60  
cctgtggcct gcaggatttg cgagatttat agggaggacg ctgggacccc caaaagctgg 120  
gaaatgggac tattggcatt cagggatgtg gctctagaat tctctccaga ggagtgggaa 180  
tgccctggacc cagctcagcg gagtttgtat agggatgtga tgtagagaa ctacagaaac 240  
ctgatctccc ttggctcttg tatgtctaag ccagaactga tcatctgtct ggaggcaagg 300  
aaagagccct ggaacgtgaa cacagagaag acagccaaac actcagcttt gtcttcttat 360  
cttactgaag acattttgcc agagcagggc ctgcaagttt cattccaaaa agtgatgctg 420  
agaagatatg aaagatgttg tcttgagaaa ttacgcttaa ggaatgactg ggaaat 476

<210> 91  
<211> 913  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1091617.1:2000SEP08

<220>  
<221> unsure  
<222> 59, 170  
<223> a, t, c, g, or other

<400> 91  
aaatccttct ttgtgaagtc caactacgct gaacatcaga gaactcacac aggagaaana 60  
ccatatgaat gtaatgaacg tggaaagtc ttctgccaga aatcagccct cacagtacat 120  
cggagaactc acatgggaga gaaactatat aaatgcaatg aatgtgagan aaccttctgt 180  
gtgaaatcaa acccttactca acattgaagg actcacacac ggtgagaaac cctataaatg 240  
taatgagtgc tggagatctt tctatgtgaa atccaacctt gtctgtgcatc agagaaatca 300  
aggagaaaaa tcctacagat gtcctgagtg tgggaaaacc ttctatgaaa agtcagccct 360  
cacaaaacat gagcgaattc acacagggga aaaaccctat gaatgtaatg aatgtagaaa 420  
aaccttcagc cagaggtcag cctcaccac acatcaaagg aaaacacaca agaagaaaac 480  
tattatcaac accctccatg tgcagaaacc tgcattctca aggc aaattt gtcaaacatc 540  
agcaaaatca tagggcagaa acactaaggt attaatatcat gtaggaaagt atttgtttgt 600  
aggttataac gaataggaat tagaaaatta acataggagt gaaaccacat gaatgctttg 660  
aacatgtgaa agctttcagc aagtattcag agctgtggag gaaaaattaa ttttaaggta 720  
tatgaatgca aaatgttctt gtgcataaaa tgtttaggta gaagtcattt ttcagatgta 780  
atagcagagt aattaaagaa aagattagaa aatgacttgt tcttaataga atgtgaaaca 840  
tgatcctttt gtaaaactat gaattattta agtcattgac acgatagcgt tgaatactgt 900  
gtgaagattt cct 913

<210> 92  
<211> 827  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1082344.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 48

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 92

```

ggagaagcca tccatcaaat gccagatctt accctgcaca agaaagtntc tgctggagaa 60
aaaccatatg aatgcaccaa gtgcaggaca gtcttcacgc atctttcttc tcttaaaagg 120
cacgtcaagt ctcaactgtg acgaaaagca cctccagggt aggaatgtaa gcaggcctgc 180
atttgtccct cacacctaca cagtcacgga agaaccgaca ctgaggagaa gccgtataag 240
tgtcaagcat gtgggcaaac tttccaacat cctcgttacc tctcccacca cgtaaagact 300
cacacagcag agaaaaccta caaatgcgag cagtgtcgga tggcgtttaa tgggttcgca 360
agcttcacta gacatgtgag aactcacaca aaagacaggc catataaatg tcaggaatgt 420
gggagagcct tcattttatcc ctgcagatct caaagacaca tgacaacaca cactggagag 480
aagccctata aatgtcagca ctgtgggaaa gccttcactt acccccaggg ttttcaaaga 540
catgagaaga cgcacacggg agagaaacct cacaaatgta aacaatgtgg gatgtccttc 600
aagtggcact cctccttccg gaaccatctg aggatgcaca caggacagaa atcccacgaa 660
tgtcagtcac actcaaaagc cttcagttgc caagtcattc tttctaaaac cagtgaagac 720
acacactaaa gagaaattct ataactttta tggggtaacc tcacattaat tcatgtataa 780
tgctccagaa aattcacacc aggagagaaa tcttaccagt atgatat 827

```

&lt;210&gt; 93

&lt;211&gt; 594

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1166249.1:2000SEP08

&lt;400&gt; 93

```

gggagctcca ggtctcgctc tcaactactct gtgtcttctg cttttagggg cgcactctgt 60
ggccctgtga cctgccccct ggaagcctag aaatgggact gttgacattt agggatgtgg 120
ccatagaatt ctctctggag gagtggcagt gcctggacac tgcacagaag aatttatata 180
ggaatgtgat gttagagAAC tacagaaacc tggccttcct gggatattgt gtctctaagc 240
cagacctcat catctgtctg gagaaagaaa aagagccctg gaatatgaag cgagatgaga 300
tggtggatga acccccaggt aggtgagagt gaatacaaca gatgacctgg atgagaggtc 360
caacgtcaag aagaaagcca gtctttaaag tgatttggca aagctgcatt ccaaaggaaa 420
ccgtttctgg aaagcctgaa atttcaacaa caaatatact ctcagatagg ggcattcttc 480
gcttatgctt ataaaatctc taagaattct actctccctt caatgatctt ccttaacgtt 540
tacagtgaca gccaaagtac tgttcattga tataaaagag tgtacaatct gact 594

```

&lt;210&gt; 94

&lt;211&gt; 501

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:799675.1:2000SEP08

&lt;400&gt; 94

```

gatttggatg agtttgattt cttttgctat ttcattgctc tagctaggac agccagtatt 60
gattgaatag aaggggtgag agcattcttg catcatgtga gatcctacag gaaaagcatt 120
ccattttccc tgactggtta tttctgctgt ggtcatttca tggatggctt ttgtattgtt 180
gaggattgat ttccagagac ttatgttact tgaggaagac accaagaagg caaagaaaag 240
gagccaggga tggctcttcc tcagggatac ttgactttca gagatgtggc tatagagttc 300
tctttgctgg agtggaaacg cctggaccct gcacagaacg ctttatacag ggctgtgatg 360
tgggagaact acaggaacct ggagtctgtg ggtgaggaaa atgttcctcc agaactgaag 420
aatctgtcct tgtgtatctt ggcttttctt ggttttgtat ctcttttgtg attttgcccc 480
atacgtgctt ttgatgtaca c

```

&lt;210&gt; 95

<211> 1655  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1178899.1:2000SEP08

<220>  
<221> unsure  
<222> 577  
<223> a, t, c, g, or other

<400> 95  
gcgcctgtga caagtaaggt tggattttct cttccctgag gtgaaggatg cccggaggcc 60  
tcggcaggac cgcgcggaaa cgggccttct gcccaaaaga tgctgcttct ctccttattc 120  
tttccctca gaatctcgct gtctccttcc aaccacctgt ggctcgcatc cccgcgttgt 180  
cactgcgacg ctgaggcgag cgagggtggc ggaagcacc gcggggcggg gagggaccct 240  
gcggggcggg actccacacc aagcctctgc tcagcgtcac ccgcttgct gtgtcctcgc 300  
aggctgcagc ttcattggcct gatgccttca ggaagtattt tgaagtcac gtggctgtgg 360  
attgggggat ttcttgtttc cactgacctg tgaggccgcg cacgtggagg gaggcacccc 420  
gggtcctcgc gcaactgtccg gcctcgcctg tgtccctagt agcagtgggc atttccagac 480  
ggtgcagctt gtggctaaag tgacaggaag atgtaggagc ttccagtctt ggatgaggat 540  
tcgaactgaa gggcttaggc ccagctgtct tggagcnaaa catctgttgt ggggtgagaa 600  
gggaaaggag ggggtgaggt gcaggggccc tctaggagg acatcggacc tgttgggccc 660  
agacctcatt agagatctag agcacggatg ggaggagcag aggcctcgatt tgtgtcctgt 720  
tcagactccc acaggctggg tgatgggtggg agatttgtgc ccattcagat gtggcggcag 780  
aggagggcaa ggccgaagga gcagacagca ccgcttcttg gggagtgttg aaggcatcat 840  
cgggagggcc gagcttagca gccaaagtga ggacagcacc ctccatgcct ggattcgtta 900  
ctcgctcggt ctcgatgttg agctgctggc atattgcagc acaactagag atgtacggat 960  
gcccccatct tgatcttaca gaatcagagg tgcagccgca agaaagctac acttctcggg 1020  
ttctctgtct tgccaacact agtggaatat gatcacatcc cagggatcag tgtcatttag 1080  
ggatgtgact atgggcttca ctcaagagga gtggcatcat ctggaccctg ctccaggagc 1140  
cctgtagagg aatgtgatgc tggagaacta cagccacctt gtctcagtag ggtattgcac 1200  
tcctaaacca gaagtgatcc tcaaattgga gacaggcaag gagccatgga tattagagga 1260  
aaaatttcga agccagagtc atctgggtga gttagtgcga gatggaattt aaagaattaa 1320  
ttaataccag tagaaactat tcaagaatga agttcaatga gtttaacaaa ggtggaaaat 1380  
gtttctgtga tgaaagcatg aaataattca ttttgaagag gaaccttctg aatataataa 1440  
caatgggaac agcttctggc tgaatgaaga cctcatttgg catcagaaga ttaaaaattg 1500  
ggagcaacct tttgaatata atgaatgtgg gaaagctttc cctgagaatt cactcttctc 1560  
tgtacataag agagcttaca caggacagaa aacatgcaaa tatactgaac atgggaaaac 1620  
ctgttatatg tcatttttta ttactcatca gcaaa 1655

<210> 96  
<211> 632  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1169241.1:2000SEP08

<400> 96  
gggagactag acctgaaatg aaagagttgg atccaaagaa tgacatttct gaagacaagc 60  
tctccgttgt tggggaggcc acggggggac ccacaggagaa tgggtccagg ggtcctggct 120  
cagaaggagt gtgggaacca ggcagctggc cagagaggcc gcggggagat gcagggtgag 180  
agtgggagcc attgggaatt cccaggggga acaaaactctt agggggctca gtacccgcac 240  
gtcatgaact gaaggcattt gccaaaccaag gctgtgtcct ggtcccacca cggctggacg 300  
acccacacaga aaagggggcc tgtccacccg taaggcgttg caagaacttc tccagcactt 360  
cagacctcag taagccccc atgcccctgc aggagaagaa aacctacgac tgcagcgagt 420  
gtggcaaggc ctttagccga agctcgtccc tgataaagca ccaaaggatc cacacgggag 480  
aaaagccgtt taagtgtgac acctgtggga agcattcatc gagcgctcgt ccctcaccat 540  
ccaccggcgc gtgcacacgg gcgagaagcc tcatgcgtgc gccagtgcg gcaaggcctt 600



cagccagcgc tccaacctac tgagccaccg gc

632

&lt;210&gt; 97

&lt;211&gt; 828

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1180090.1:2000SEP08

&lt;400&gt; 97

```
acctgcttga gccagcgacc gatcctttgc tgctctgctc tgccggcttc ttgaggagcg 60
tcctggggag ttggaattca gaaggaatct ctcagacgca aatctccatc tcctccctga 120
cctcacctgg gcagcctttt cagactgtct ctgccagaa tgtcaggaga ggccacagtc 180
ttggcctacc atgctccaga agaacaggaa ggacttctag ttgtcaaggt tgaagaagaa 240
aattatgttt tggaccagga ctttggcctt caggaaaacc cctggagcca agaggtattc 300
cggcagaagt tcaggcagtt tagttactct gactccactg gccctcggga ggctctgagc 360
cggtctcgag agctttgctg tcagtgggtg aggcgggagg tgcactcaa ggagcagatc 420
ctggagctgc tgatgttgga gcagttcctg gccatccttc ctgaggagct gcaagcttgg 480
ctgcgagagc atcggccaga gaatggagag gaagctgtga ctatgctgga ggagctggaa 540
aaagaactgg aggagccaag gcaacaggac acaactcatg gccaaagaaat gttctggcag 600
gaaatgacat ccacaggagc actgaagtct ctgtctctga atagcccggg gcagccctta 660
gagaaccagt gcaagactga gactcaggag tcccaggctt tccaggagag aggtgagaag 720
ccccagtcca tgtgggggtg aggatggagg tgagagctat caactgtggc acggttcatt 780
ttcttaacga cagccacttg tggctatcac taagtattgt ggaaaggc 828
```

&lt;210&gt; 98

&lt;211&gt; 528

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2049322.1:2000SEP08

&lt;400&gt; 98

```
gcggatcgcg tggagtgaag gtcacgccgc ggccggattg acttctaaag agtcatgctg 60
tgtgatgaaa aagcccagaa gagaaggaag aggaaagcaa aggagtcagg gatggctctt 120
cctcaggagc acttaacatt cagggacgtg gccatagaat tctctcaggc ggagtggaaa 180
tgectggacc ctgcacagag ggctttatac aaggatgtga tgttggagaa ctacaggaac 240
ctggctctcc tgggaatctc tcttctgac ctgaatatta actccatgtt ggagcaaagg 300
agggagccct ggtctggtga gagtgaagtg aaaatagcaa aaaattcaga tggggaggag 360
tgcataaaag gtgtgaacac agggagcagc tatgcattgg gaagcaatgc agaagacaaa 420
ccaattaaaa aacaacttgg agtatccttt cacttacatc tgtctgaact ggagctattt 480
ccagatgaaa gggtaataaa tggatgtaat caagttgaaa actttatc 528
```

&lt;210&gt; 99

&lt;211&gt; 486

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:809074.1:2000SEP08

&lt;400&gt; 99

```
ggcgggttgc ttgtctctcc ttccgccggag ccggctcttc ctgttagtct ccgctgctag 60
ttcttggctc tgggaggccc aggtggctct gcagcagcct ctgccaccct gtgacctgcy 120
agtattggga catccctagc tgacgccagg acaccgggga agccaggagaa tggactcggg 180
ggcttttgag gatgtggctg tgaactttac ccaggaggaa tgggctttgc tagattcttc 240
tcagaagaat ctctacagag aagtgatgca ggaaacctgc aggaacctgg cttctgtagg 300
aagccaatgg aaagaccaga atattgaaga tcacttcgaa aaacctggga aagatataag 360
```

```

aatcatatc gtacagagac tgtgtgaaag taaagaagat ggtcagtatg gagaagttgt 420
cagccaaatt ccaaatcttg atctgaacga gaacatttct actggattaa aaccatgtga 480
atgcag                                     486

```

```

<210> 100
<211> 756
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:805158.1:2000SEP08

```

```

<220>
<221> unsure
<222> 17
<223> a, t, c, g, or other

```

```

<400> 100
tactgaatga gcatggnaaa actcacacga gggagaagcc ctttgcattgt gtggttttgtg 60
gaaaaatatt tagaaattcc tcatgcctta ataatacatgt tcgaattcac actggaataa 120
aaccctatac atgcagctac tgtgggaagg ccttcaactgt gcgctgtggc cttactagac 180
acgtacgaac acacacgggc gagaagccat acgcgtgtaa ggactgcggg aaagccttct 240
gtacatcctc gggcctttac tgagcatgta aggactcaca ctggagagat aaccatata 300
tatgtaaaga ttatgggaaa tccttcaactg tttcttcaag cctgactgag cacgcgagaa 360
tccatatcac ggacgagaaa ccctacgaat gtaagcagtgt tggcaaagcc ttcacagggc 420
gctcaggctc actaaacaca tgcggacaca caccggggag aagccctatg aatgtaagga 480
ctgtgggaaa gcctacaata gggtttatct actgaatgag catgtgaaaa ctcacacaga 540
ggagaagccc ttacatgta cggtatgcag gaaatccttc agaaattcct cgtgcctgaa 600
taagcacatt catattcaca ctggaataaa accttatgaa tgtaaggact gtgggaaaac 660
attcactgtt tcttcgagcc taactgagca catacgaact cacactggag agaaccttat 720
gaatgtaaag tatgcggaag gccttcacca catcct                                     756

```

```

<210> 101
<211> 892
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:1172697.1:2000SEP08

```

```

<400> 101
ctgagctcca agtctcgtct tcagcgctct gtgtcctctg ctccctagagg tccaggctct 60
gtggccctgt gacctgcagg tattgggaga tctacagcta agacgccagg aacccttgga 120
agcctagaaa tggagaacct gaagtctgga gtgtatcctc tcaaggaagc aagtggatgc 180
cctggggctg acaggaatct tctggtgtac tctttttatg aaaaggggccc attgacattt 240
agggatgtgg ccatagaatt ttctctggag gagtggcaat gcctggacac tgctcagcag 300
gatttgtata gaaaagtgat gttagagaac tacagaaacc tggcttctct ggcagggtatt 360
gctgtttcta agccagacct gatcacctgt cttagagcaag gaaaagagcc ctggaatatg 420
aagagacatg cgttggtaga tcaaccccca gttacatatt ctcatcttgc ccaagacctt 480
tgccagagac agggcataaa agattctttc caagaagtca tattgagaag atatggaaaa 540
tgtggacatg aagattttaca gtttaagaaca ggctgtaaaa gtgtggatga gtgtaactctg 600
cacaagaat gttatgatga actaaaccag tgtttgacaa ctaccagag tgaaatattt 660
caatatgata aatatgtgaa tgtcttttat aaattttcaa atccaaatat acaaaagata 720
agacatactg gaaagaagcc tttcaaatgt aaaaaatgtg acaaatcggt ttgcatgctt 780
ttacacctaa ctcaacataa aagaattcat attagggaaa attcttacca atgtgaagaa 840
tgtgtgtaag tctttaactg gttctcaacc cttactagac acagaagcgc gg                                     892

```

```

<210> 102
<211> 670
<212> DNA
<213> Homo sapiens

```

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1174107.2:2000SEP08

<220>  
<221> unsure  
<222> 663  
<223> a, t, c, g, or other

<400> 102  
gggcggttct tcaactgctct gtgtcctcag cgtgtgtggc ttcgtgacct gaagatactg 60  
ggaatccat agctaagatg ccaggacccc ctgaaagcct agacatgggg ccgttgacat 120  
ttagggatgt ggccatagaa ttctctctgg aggagtggca atgcctggac actgctcagc 180  
aggatttgta taggaaagtg atgttagaga actacagaaa cctgggtcttc ttggcaggta 240  
ttgctgtctc taagccagat ctggtcacct gtctggagca aggaaaagat ccctggaata 300  
tgaagggaca cagtacggta gtcaaacccc cagggttttct taccgccatc tgtgacagct 360  
tcttgatctg tcccaagtta tatgttctca ttttgctgaa gacttttgcc cagggccagg 420  
cattaaagat tcttttcaaa aagtgatact gagagaatat gtaaaatgtg gacataagga 480  
tttacagtta agaaaaggat gtaaaagtat gaatgagtgt aatgtgcaca aagaaggta 540  
taatgaacta aaccagtatt tgacaactac ccagagcaaa atatttcaat gtgataaata 600  
tgtgaaagtc tttcataaac ttttaaattc aaatagacat aacacaaaac atactggaaa 660  
ganaccttca 670

<210> 103  
<211> 1285  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1177434.2:2000SEP08

<400> 103  
ggcactttgg cttgtgtcag ttccatccgc ggggtgccga tctggacctg ggtgctgaca 60  
gcgagaaggc gcgaggagag tcgttttctc agctgcacag ccggggcctg acggtcgccg 120  
gcggtggtga cagctttgct cttgtctccg cccggatcgt ccaccgctcc cggcccgctc 180  
cgcccagagt ccgatggcgg cggcactgag ggccccgacc cagattgccg ctagtactta 240  
caggctccgta agaagcactt gaagacccta caaggcaggt ttttgtagcc tttgaggatg 300  
tggccattta cttctccag gaggagtggg agctccttga tgagatgcag aggcctcctgt 360  
accgcatgt gatgctggag aactttgcag ttatggcatc cctaggttgt tggcatggaa 420  
tggaggatga agagatacct tttgagcaga gcttttctat aggaatgtca cagatcagga 480  
ttcccaaggg aggtccttct actcagaagg cttacccttg tgggacatgt ggcctggtct 540  
taaaagacat tttgcacttg gctgagcacc aggaaacaca ccagggcaga aaccatacat 600  
gtgtgtgctg tgtgggaaac agttctggtt cagtgcacac cttcaccagc accagaagca 660  
gcacagtga gagaaccct ttagaagtga taagagcagg ccctttcttc tgaacaactg 720  
tgctgtgcaa tcattggaga tgtcttttgt gacaggggag gcttgtaagg acttcctagc 780  
cagctcaagc attttcgagc accatgcccc tcacaatgag tggagccac acagcaacac 840  
caagtgtgag gaggcctctc actgtggaaa aaggcattac aaatgcagtg aatgtgggaa 900  
aaccttcagc cgcaaagact cacttgttca acaccagaga gtccacactg gagaaaggcc 960  
ttatgagtgc ggtgaatgtg ggaaaacctt tagccgcaaa ccataacttg ctgagacca 1020  
gagaatccac actggagaaa tgcccttatga gtgtggcata tgtgggaaag tttttaatca 1080  
tagctctaac cttattgtac atcaaagagt acacactgga gcaaggcctt acaagtgcag 1140  
tgaatgtggg aaagcctata gtcacaaatc tacacttggt cagcatgaga gtatccatac 1200  
tggagaaagg ccatatgagt gcagcgaatg tggaaaatac tcttggtcac aaatacagac 1260  
tcattaaaca ttggagcgtt tgcgg 1285

<210> 104  
<211> 774  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature

<223> Incyte ID No: LI:1184255.1:2000SEP08

<400> 104

```

caagaacttc aggcacaaat tttctctaac caatcatcag agaagtcaca cggcggaaaa 60
accttacaaa tgtaatgaat gtggcaaggt ctttagtctg ctttcatacc ttgcacggca 120
tcaaataatt catagtacag agaagcccta caaatgtaac gaatgtggaa gagcatttca 180
caagcgctcgg ggccttatgg cccatcttct aatccatact ggagagaaac cttacaaatg 240
taatgaatgt gacaaagtct ttggggcgcaa attataccta accaaccatc agagaattca 300
tactggagag agaccttaca agtgtaatgc atgtggcaag gtcttcaatc aaaatccaca 360
ccttttcacga catcggaaaa ttcattgctgg agagaattca ctgcgtacct tacagatgga 420
atgaatgtgg caaaatcttt agttagaatt cacacgtagc acaacattgg aggactcagg 480
agaaaaacct tacaaatatc ataaatgtgg caaagaatgt agtttgcatt gaagcctcat 540
cactcatctc ttgatccaca ctgaaaggaa accctacaaa tgtaaagtta ataacatata 600
taataatcta taaagagaga acagttatat cagaatagag catttaatga aaactaccag 660
tatagcaata ttcttgagta ggattcacat ttaactgtgg caccgtaatt gtaactcttg 720
atttagaatg tacatacttc tcatgtttta agtaataaag tttaaagttt tttta 774

```

<210> 105

<211> 487

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1164555.1:2000SEP08

<400> 105

```

cacacaattc caggaaatat gcatgagaga aaacccttgc caaagtaatg agtgtggcaa 60
agccctttta ttatagctca ctcttaagga gacaccacat aaccatttca agagagagag 120
aatataaatg tgatgtatgt ggcaagatct ttaatcagaa gcaatacatt gtatatcatc 180
atagatgtca cactggtgag aaaacttaca agtgtaatga gtgtgggaag accttcactc 240
agatgtcatc ccttgtatgc catcgtagac ttcatactgg agagaaacct tacaagtgtg 300
atgagtgtgg caagaccttc agtgagaagt catcccttag atgccatcgt agacttcata 360
ctggagagaa accttacaag tgtaatgagt gtggcaagac ttttggtcga aattcagccc 420
ttgtaattca taaggcaagg tcttcactct gtggtcattc catgcagaac atcagaaaat 480
tcattttt
487

```

<210> 106

<211> 1320

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:238666.4:2000SEP08

<220>

<221> unsure

<222> 1196, 1276, 1281

<223> a, t, c, g, or other

<400> 106

```

caggcgcaact tccagttctg cccacccgct gcggccattg tccgaccccg gtgcggctga 60
ggccccctttg ggcagccctt ccgcagatca gaattggaga caaccgagcc ttcggcgggg 120
gcgggaggag ctgcccagag ctctgggttg gccggaggtc gcgaaatccg gagcccccca 180
gaggcggtga ttctgactcc tgcccgggag ccgggccctg gcgaggcagg aatggccccg 240
aggcctccga ccgcgcgcgc ccaggaatca gtgacattca aagatgtgtc tgtggacttc 300
accaggaag aatggtacca tgtcgaccct gtcagagga gcttatacag ggatgtgatg 360
ctggagaact atagccacct ggtttctctt ggatatcaag tttccaagcc agagggtgatc 420
ttcaaattgg agcaaggaga agagccatgg atatcagagg gagaaatcca acgaccttcc 480
tatccagact ggaagaccag gcctgaagtc aaatcatcac atttgagca ggatgtatca 540
gaagtatccc actgcacaca tgatctctta catgctacat tagaagactc ctgggatgtt 600
agcagccagt tagacaggca acaggaaaac tggaagagac atctgggatc agaggcatcc 660

```

accagaaga aaataattac accacaagaa aattttgagc aaaataaatt tggtgaaaat 720  
tctagattga acaccaatatt ggttacacaa ctgaacattc ctgcaagaat aaggcctagt 780  
gaatgtgaga cccttggaag caatttgga cataatgcag acttacttaa tgagaataat 840  
attcttgcaa aaaaagaaacc ctataagtgt gataaatgta gaaaagcctt tattcataga 900  
tcatcgctta ctaaacatga gaaaacacat aaaggagagg gagctttccc taatggaaca 960  
gatcaaggaa tttatcctgg aaagaaacac catgaatgta ccgactgtgg gaaaaccttt 1020  
ctctggaaga cacagcttac tgagcatcag agaattcaca ctggggagaa accctttgaa 1080  
tgcaatgtat gtggaaaggc cttcaggcat agctcatctc ttggtcagca tgagaatgct 1140  
cataccggag agaaacccta tcagtgtagt ctctgtggga aagccttcca gcgcantcct 1200  
cccttggtta acaccagcga attcacactg gagagaaacc ctatcgatgt aaatctatgt 1260  
gggaggtcct ttaggnatgg nacatccctc actcaacacg agggtcacac acagtggaga 1320

<210> 107

<211> 581

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1166752.1:2000SEP08

<400> 107

ctcgtatgtg tagtctgatg ctgtgtgaga tgagtcattc ggctgaaggt tttcctgcat 60  
tccttgcatc tgtaggggtt ctaccagtg tgaattctct catgctaggt aagggtgtgc 120  
ctctgggtga ggggttttaac acattcggtg catttatagg gtctctctcc cgtgtggatt 180  
ctttgatgct gaatgagagt tgattgacga gtgaagtatt tcccacaatc taaacactga 240  
ttgcatttct tttgtgtatg atttctctta tggagagtaa gctgcatatt gtggcagtat 300  
agatttccac attccttaca tacgtttctt gtacatttta tcaactgtgtc aaactgcat 360  
aatacttcat taatagccct ctttgcagat aactccttaa tctctgcacc tggcacatct 420  
gaaaggggaa aaaaaatcaa tgaacagaag caaattaaac taatgtaata tacctatctt 480  
ttatcttttc aatcccaact taggcaccta ctcaagaaac taagatacca gtttttaagg 540  
ttatattgat aagaatgctt attgaaatag gctttatatt g 581

<210> 108

<211> 577

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:2049654.1:2000SEP08

<400> 108

gcgatgtggc gcttgcgac tctcgccgcc ggcagaggct cctcgaagag cgacacgggg 60  
ctgaccaggc acgggtggtca aagccgcaga gggagagcgg gagcggtcgt gagatcgtct 120  
ggggagaagg gcggaggcaa agccgaggag gtgcgggttg tggccattc tggaggacgc 180  
tgatcgaatg ccccaaaactt cccggaatgt gtgtggaccc ttctagcctg aggagtcctg 240  
caggtgtgaa gctccacacc tgctccata gcactttgcc tgtccctaag agggctcatc 300  
ggagaagaaa gaatggctgt cagccacctg ccaaccatgg tccaggaatc ggtgaccttc 360  
aaggatgtgg ctatactgtt caccaggaag gaggggggc agctgagccc cgcccagagg 420  
gccctgtaca gggacgtgat gctggagaac tacagcaacc tggctcact gggactctta 480  
ggacccaaac cagatacgtt ttcccagcta gaaaaaggg aagtgtggat gccagaggac 540  
accctggag gcttctgtct tgatggagtc tcaactc 577

<210> 109

<211> 472

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:242665.2:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 421, 460

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 109

gaaatagtac	agaagagcaa	agaataataa	tgagtcagct	agacagaaga	caggagtcac	60
cagattttgta	ccctagttac	tcgacccatc	aactgactgt	ggtttttggg	tttgtctgct	120
tcttggttatt	gggcctcgag	aaaatgcagc	cagacctggc	ttgggtccag	gaacaataag	180
aataaagact	tgctgggttc	ccaccacaca	tggtgtcgtg	agctcactac	acaccctgtc	240
cctggagcat	tggtatgcag	cctggagcca	ggaagatgac	agtgttgtgc	acagctctgc	300
actggtaatg	cactatgcat	gagtcgccct	caggaccttg	tgttactgct	gctggttggc	360
tggtccaggca	ctgctggata	cagcaagggc	actggcgata	caggcagcag	cttcagttca	420
nacctgccaa	ctttcagcct	aagccaaggg	gcaccacatn	cagcacgtcc	ac	472

&lt;210&gt; 110

&lt;211&gt; 5213

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:208637.1:2000SEP08

&lt;400&gt; 110

cggtccgcgcg	tcgacgtcga	cccatactgg	agcgacgttt	aaagaagggg	cagaatcgct	60
ggggagtgcg	gcttcttctt	gttgggggac	tcccagcctt	ccgcgcgtcc	ggaggaggag	120
aagcggcgccg	gccgggaagc	aggcatggag	agtacaaaac	tgatttctgc	tacagacatt	180
cagtactctg	gcagtctgct	gaactccttg	aatgagcaac	gtggccatgg	actcttctgt	240
gatgttaccg	ttattgtgga	agaccgaaaa	ttccgggctc	acaagaatat	tctttcagct	300
tctagtacct	acttccatca	gctcttctct	gttgctgggc	aagttgttga	actgagcttt	360
ataagagcag	agatctttgc	agaaattctc	aattatatct	atagttctaa	aattgttcgt	420
gttagatcag	atttgcttga	tgagttaatt	aaatcagggc	agttattagg	agtgaattt	480
atagcagagc	ttgggtgtccc	attgtcacag	gttaaaagca	tctcaggtac	agcgcaggat	540
ggtaataactg	agcctttacc	tcctgattct	ggtgacaaga	accttgtaat	acagaaatca	600
aaagggtgaag	cccaagataa	tggtggctact	ataatgccta	ttataacaga	gtctttttca	660
ttatctgccg	aagattatga	aatgaaaaag	atcattgtta	ccgattctga	tgatgatgat	720
gatgatgtca	ttttttgctc	cgagattctg	cccacaaagg	agactttgcc	gagtaataac	780
acagtggcac	aggtccaatc	taaccacagg	cctgttgcta	tttcagatgt	tgacacctag	840
gctagcaata	actcgcccc	tttaacaaat	atcacaccta	ctcagaaact	tcctactcct	900
gtgaatcagg	caactttgag	ccaaacacaa	ggaagtgaag	aattgttggt	atcttcagct	960
ccaacacatc	tgactcccaa	tattattttg	ttaaatcaga	caccactttc	tacaccacca	1020
aatgtcagtt	cttcacttcc	aaatcatatg	ccctcttcaa	tcaatttact	tgtgcagaat	1080
cagcagacac	caaacagtgc	tattttaaca	ggaaacaagg	ccaatgaaga	ggaggaggag	1140
gaaataatag	atgatgatga	tgacactatt	agctccagtc	ctgactcggc	cgtcagtaat	1200
acatcttttg	tcccacaggc	tgatacctcc	caaaatacca	gttttgatgg	atcattaata	1260
cagaagatgc	agattcctac	acttcttcaa	gaaccacttt	ccaattcctt	aaaaatttca	1320
gatataatta	ctagaaatac	taatgatcca	ggcgtaggat	caaaacatct	aatggagggt	1380
cagaagatca	ttacttttaga	tacagctact	gaaattgaag	gcttatcgac	tggttgcaag	1440
gtttatgcaa	atatcggtga	agaatactta	tgacatagtg	attccttgtc	aaagatgacc	1500
ctgatgaagg	ggaggccaga	cttgagaatg	aaataccaaa	aacgtctggc	agcgagatgg	1560
caaaacaaac	gtatgaaagt	aaaacatgat	gatcactatg	agttaatagt	agatggaagg	1620
gtctattata	tctgtattgt	atgcaaaagg	tcatatgttc	tgtctgacaa	gcttgcgagg	1680
acaftttaac	attcattctt	gggagaagaa	gtatccgtgc	cgttacttgt	gagaagggtat	1740
cttctctctg	cagaaatata	gcacaaagca	tgaaaattca	tccacacagg	gggagcgagg	1800
gtatcagggtg	tttggcctgt	gggcaaatct	ttcatcaact	atcagtttat	gtcttcacat	1860
ataaagtcag	ttcatagtca	agatccttct	ggggactcaa	agctttatcg	tttacatcca	1920
tgccaggtctt	tacaaatcag	acaatatgca	tatctttccg	atagatcaag	cactattcct	1980
gcaatgaagg	atgcatggta	tcggtgtata	cggttgacac	tggaacagaa	cctccagtag	2040
ggaccactac	atctactcag	aacaagccaa	tgacctggga	agatattttt	attcagcagg	2100
aaaatgattc	aatttttaaa	caaaatgtaa	cagatggcag	tactgagttt	gaatttataa	2160
taccagagtc	ttactaaact	cctttgaaat	actagaaagt	tttgtttttg	atgatggggc	2220
aggggttttca	gaagatctgt	aaaacaaatt	aaggtgcgaa	caagttaatt	tgatctgcca	2280

```

cattatctga aggaagtgtg gtgggatttt tgttgataat ttttagaagc aaattttcct 2340
gaaagttttg agtagagggt agacccccct ccccaagtat ctgtttatat agtttagttt 2400
cagctcattt aaaagaggca aacaattaaa agcttggaga gatagtttcc tgaatagaat 2460
ttgaagcagt ctgaatgttc tttgaaaata actggagtta ttagcatacc ctagtacatc 2520
ttacagcttt ccccttccat gttagcactt tactgctgaa ttctcaattt tcttaacatt 2580
gagacaataa atgtgtgttt tgtcttgtat atggcataaa gagtaataa gtttttagagt 2640
tgttctggaa aatgtcagaa taagtcagta cttgggttgt gtaatctgct agtccaagcg 2700
aacagcaacc tcctgctacc ctccctctat gaaaatagcc atgcagacaa gtctctcatc 2760
tgaagaacac atatagatat tagactcaat tagaattact cctggctttc attgccatag 2820
atctgtaaaa gacttttggtg gctagaccac tttatacctt ttgcagtgtg gtctctgggg 2880
gcaaaaaact aatgaaaaca atctctgtaa tggcagatag gaggagatga aaaagttctg 2940
ttgcatggat ttttaattct ctggctacca catagtagag aatggaatga agatttcctt 3000
ttggcttctt aagggttaaaa atattcccat gaacatgaaa attttcaa atttgaattc 3060
aaagccacca aatgtatctt tatgtataaa tccttgtaaa tgatagattc catgggtgag 3120
actttacata tttttggtgg gaggtactg gcataatatt ttaaattgtt agtattgcgt 3180
agaatctcca ctagagaagt ctttatttga aatagttgaa tcagtgatct agtattttcc 3240
tttcggcaag atttgttagg tttttacccc ttctcaaata agttttattc catctgcaaa 3300
ttgctgcaat attatagtaa tcagaaacta cataaggaat gttatatagg cttgtcagtt 3360
cccgtttttc ttgacaacaa taaataccac ttttaaaaat gacacatatt taaacactta 3420
gaaaataaag ttaacactta ctgaagtgtc agtactaaac tgtgctagta ctaaaagaaa 3480
acaggatgga acatacatat agcctagcat ttataacaga attggtgaac gtctgtaaat 3540
gatttttttt ttttttggca agggaaaaat ttgttccggg aaaagtttgt tgtgcatagt 3600
tatttagtct ttgtaacctt gcttaagttt ttcttagtcc aacatagata ttttctttct 3660
cctgaccatg tattttaaaa tatagtctat ttcttgactt tgaacttaaa gctttaatca 3720
taatttctca tgtatacatc gttcttctga tggtaagctg gatttgaagg tagtggtttc 3780
agtgtttctt aagttggtag ctgagggtat caggcatcag ttcatgcaat aatacaagaa 3840
aaaaaatcct ttgcttgcca agaggtagag tgatgtgcat ttatctgttt tctgttctgt 3900
aagtctagac cttcaaacca tttgtaaact aacccctggg aaatttgaaa ttacctgata 3960
acttaagact ctgtgatctc tggaaatcac atagtgttct tttttgtgta gatattaata 4020
acattactct ttgactatag tgtgcactct gaaatgtact cagtgaatat ttcgttttga 4080
gtttcattaa tgctatttca ccagttagac ataattactt ctaccgatgt gaatgatacg 4140
ggatgccggc agagcttcca gatctttcag actcaactgc tagggcaatt agtttgtcat 4200
aataaaaaact tggcagattc tacaagtcct attatgacaa accagggaac taattctatt 4260
atggaaaaat aatccatcct gaaataataa ggtatgtaat tattttgctt ggtgggtggtg 4320
tggccctgta aattctctgg attttacctt taaactccgg ggctactaaa ggtatctata 4380
ctaggaggta ttttggggaa ggacgcagaa actggaaaat taaacttgta ttttctgcca 4440
gaagaactct taacttggca tgtgtctcag ggtcttcaag ttttctata agtttccata 4500
tccaaagttc cagaattcca tgtgaaatac ttctttgggg gcaaaagtcc ttcatctctg 4560
ggtatttatt gggattggga aatctgtagg caaggatgct gtttaaaaat taccatattg 4620
tttttttate ttatacttta gctctctggg tatttggact tccttttctt gtgtgaagtt 4680
agcttcaaat ttgctcctat gctaaattac ctgtaaatat tctggatagg aactacttga 4740
aatagtaatt tgttaaaaga tatgacaaaa tgaaaatgct taaactacag aaatttaaaa 4800
atgccataac aatcttgcca agactaactt taaagatata ctttaaatga ttgattatga 4860
tctttggtgg taacgatccc ccagcacaca aactactatg aagaaataat gccgcatttt 4920
tccccccatt gtaccaaaga agataaaaaca aatggtaaac actgattcaa ggtattttgt 4980
cattgtcaag gcatgcatat tcctaaagaa ttaaattgcta acttaacagc actggctttc 5040
tggctggtca actatatgaa accttgttca ttctccgag tactgtaatg ttcacacttg 5100
tacaatcttc cctgtcatga ctttaagttc tacttttcat taaccatggc ctgatattag 5160
ttcttagagc ttcttgtggc aaaaataaaa tgatttaatt ctgaaaagtc gac 5213

```

&lt;210&gt; 111

&lt;211&gt; 998

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2051808.1:2000SEP08

&lt;400&gt; 111

```

gcaccgacca tggcctccaa gtgtcccaag tgtgacaaga ccgtgtactt cgccgagaag 60
gtgagctccc tgggcaagga ctggcacaag ttctgcctca agtgtgagcg ctgcaacaag 120
acactgacgc ctggcggcca tgctgagcat gatgggaaac ccttctgcca caagccctgt 180

```

```

tatgccacac tgtttggacc caaaggcgtg aatatcgggg gcgctgggtc ctacatctac 240
gagaagcctc cgacggaggc ccctcagggtc actggcccca tcgagggtccc tgtggtacga 300
actgaagagc ggaagaccag cggggccgcc aagggtccca gcaaagcctc tagtgtcacc 360
acgttccactg gggagcccaa catgtgccct cgatgcaaca agagagtgtg cttcgctgag 420
aagggtgacct ccctgggcaa ggactggcac cgaccctgcc tgcgctgtga gcgctgctcc 480
aagacactga ccccaggcgg gcatgctgag catgatggcc aaccctactg ccacaagccc 540
tgctatggaa tactctttgg acccaaagga cgtgaatact ggtgctgtgg gcagctacat 600
ctacgacaag gaccgggaag gcacagttca gccttagatc tgcagatgct gccctcgcg 660
gtcccctatt tgaccctgag ggagagtgat ctgctgccta gtccctgcctc atgtgtgtct 720
cgcttgaat tccagcacct aagtgtgga ggagaaagcc cagctagtcc tagaatttca 780
gctcccttgg tcaccttggc atgtcctatg ttgcccgtgg tttactttct gtctgtgtgt 840
ctccatagca ccattgggtcc tgtgtcccca tgtccctgat aactctocaa ggtggctgtc 900
ctatgatatc cctttgcccc cactcgcccc ccagtattat ttatgctctg cttgccgggtg 960
atgggcatga gctagctcac agcattcccc gggtagacg 998

```

&lt;210&gt; 112

&lt;211&gt; 470

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1175136.1:2000SEP08

&lt;400&gt; 112

```

ccaccatctt tctgggtccg ggaggatgga gtcacggggc tcccaatgcc gtggggattg 60
tttacacggc cgggttaacc cggggaatta gagaaactcca tcgtgctgac agcatccttt 120
tattcacgcg ctccgtaatt tgcaaagagg aggaagggag ggacttcttg gcttctccca 180
gcatagcccc agtttatgcc atctcagaac tatgaccttc cccagaaga agcaggagaa 240
taatgaccaa gtttcaggag gctgtgacat tcacggatgt ggctgtgggtc ttctccaggg 300
aggaactgcg attgctcgat cttacccaag aggaagctgt accgagatgt catggtgggg 360
aactttcaag aacctgggtt gcagtgggca gcaagcatca aaataatgat ggaaacaccc 420
cacaaaattg gcaataaata acctttcaaa tcaaaaaaaa caaaaaaggg 470

```

&lt;210&gt; 113

&lt;211&gt; 1843

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1177337.1:2000SEP08

&lt;400&gt; 113

```

ggcaaagcct ttaatgagtc ctcaaacctt actacacata agatgattca tactggagag 60
aaaccctaca aatgtgaaga atgtggcaaa gcttttaacc ggtccccaca acttactgca 120
cataagataa ttcatactgg agagaaacct tacaaatgtg aagaatgtgg caaagctttt 180
agccagtcac caatccttac tacacataag agaattcaca ctggagagaa accctacaaa 240
tgtgaagaat gtggcaaagc ttttaaccga tcctcaaate ttactaaaca taagataatt 300
catacaggag agaaatctta caaatgtgaa gaatgtggta aagcctttaa ccaatcctca 360
actcttacta aacataggaa aattcatact agacagaaac cctacaactg tgaagaatgt 420
gacaatacat ttaaccagtc ctcaaacctt attaaacaaa ataattcata ctggagagaa 480
actctacaaa tgtcaagaat gtgggaaagc ctttaagcag tcctcaactc ttactaagca 540
ttaaatattg gccgggtgcg gtggtttatg caaaatggct cccagcattt tggggaggctg 600
aggtgggttg atcacaaggt caagagatcg agaccatcct ggccaacatg gtgaaaccct 660
gtctctacta aaaatacaaa aattatctgg gtgtgggtgg acgtgcctgt attccatct 720
actcgggagg cttaggcagg ataatcactt gaacctggga ggtggaagtt gcagtggacc 780
aagattgtac cactgcactc cagtttggca acagagttag actccgtctc aaaaaaatt 840
tatactgtac aaaaacccca caagtgtgaa aaacatggca aagcctttaa gaagccctca 900
attcttaaca gacataagat aatttatact ggagagaaac tctacaaatc agaaagatgt 960
gactactttt gacaacgcct caaacttttc taaccataaa agtaattata ctgggtgagaa 1020
atcctagaaa tctgaagagt gagataaagc ctttaaatgg ttgtcacact tcattgtagg 1080
taagataatt catactggag aaaacgccta catgtgtgaa caatatggca aaacttaatt 1140

```



```

ctcacacttt attgctagga aagcatttat acttgagata aattatacaa atataaagac 1200
tgtgaaaaag ccatcattat ctgctcacat tttactcaac accagagagt tcctgcttaa 1260
taaaagcatt ataagtgc aa ttactgtcaa aagatctttc agaaaatata agctttttta 1320
gtgaagagta tttattttga agttgaacat cacaaatata aaaagggttg tagtgccctt 1380
acttgatca cagatctttt tgtagacatt ttcttctaga ggaaaactct gaagcagttg 1440
atcaaacttt gttcaacatc cagggaaatt atattgaaaa attgtgcaaa tataataaat 1500
ttgaaaaaac atttttcaaa aactacaaat tataaaacac caaagagttt atactaaaat 1560
atatttttgc aggtgcagta aatatgaaag atatttaatc caaaattgtc tatgtaaata 1620
tcagaataat ccacagtaga aatatctagg gccactcaat gtttagacac tactgtaaat 1680
cagagtgtcg cgtataaaaa acaatacaaa actaaagata gcataaacat tatttgtata 1740
taactttaaa agaagtagaa tatttttttg agagttataa ttacattcaa attatacttt 1800
ttatgaaaaa acagggtttt ttcgaagtga ataataatga agt 1843

```

&lt;210&gt; 114

&lt;211&gt; 1495

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1165056.1:2000SEP08

&lt;400&gt; 114

```

gcagcgcgct agtttccctt tgttttagatt cagtctgggc ttcccagctc ccccgcgctt 60
ctgtacccgg gatctgagag tcaacacaga ccttgaaatc cccgcaccgc tcctccacc 120
ccgtgtaaat tcaggcgtct ccgtgagagt ccggcgctcg ctccctgtg tgtaaataatc 180
gctcggcgac gggtcctgtc cccgctcggt ctgccttggg ccagacacag gagtgttgcc 240
atgttgtcca aactgggtct gaaattctgg agtcaagcag tctgcctgcc tcgttctcca 300
aacgtgtctg gattacaggc atgagccact acgctgggc aagtattcgg tatttataaa 360
atctacagta gtacacagtt atgtcgtagg ccatcacatt cggtcaccac acactcactg 420
actcatccag agcaatttct agtctgcaa gcgccattca tggattgact tataaacagt 480
catgtatgt gatgaagaag cccagaagag gaaagcaaa gagtcaaggga tggctcttcc 540
tcagggacgc ttgacattca tggacgtggc catcgaattc tctcaggagg agtggaaatc 600
cctggaccct ggacagaggg ctttatacag ggacgtgatg ttggagaact acaggaacct 660
ggctcttctg ggaatctgtc ttctgacct aagtattatt tccatgttga agcaaaggag 720
agagcccttg attctgcaa gtcaagttaa aatagtaaaa aatacagatg gaagggaatg 780
tgtcagaagc gtgaacacag ggaggagctg tgtattggga agcaatgcag aaaacaagcc 840
tattaaaaat caacttggat taacccttga gtcacatctg tctgaattgc agctgtttca 900
agccggaagg aaaatttaca gaagtaatca agttgaaaag tttacaaacc atcgttcctc 960
agtttcacca cttcaaaaaa tttcttctag tttcacaaca cacattttta ataatatatg 1020
aatgatctt attgatttcc cactactccc acaagaagag aaagcataca ttagaggaaa 1080
atcttatgaa tatgaatgta gtgaagatgg tgaagttttt agagtccgtg caagccttac 1140
taaccatcaa gtaatccata ctgcagagaa accttacaaa tgtactgaat gtggcaagggt 1200
cttcagtcgc aattcacacc ttgtagaaca ttggagaatt catactggac agaagcctta 1260
caaatgtagt gaatgtgaca aggtgtttta tcgcaattca aaccttgcac gacatcaaag 1320
aattcatact ggagagaagc ctcaaaaatg taacgaatgt ggcaaagctt ttagagagtg 1380
ttcaggactt actaccatc ttgtaatcca tactggagag aaaccttaca aatgtaatga 1440
gtgtggcaag aacttcaggc acaaattttc tctaaccaat cattagagaa gtcac 1495

```

&lt;210&gt; 115

&lt;211&gt; 1247

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1175250.1:2000SEP08

&lt;400&gt; 115

```

cttcgcggga gccggctctt cctgttagtc tccgctgcta gttcttggct ctgggaggcc 60
caggtggctc tgcagcagcc tctgccacc tbtgacctgc gagtattggg acatccctag 120
ctgacgccag gacacccggg aagccgagga atggactcgg tggcttttga ggatgtggct 180
gtgaacttta cccaggagga atgggctttg ctgattctt ctcagaagaa tctctacaga 240

```

```

gaagtgatgc aggaaacctg caggaacctg gcttctgtag gaagccaatg gaaagaccag 300
aatattgaag atcacttcga aaaacctggg aaagatataa gaaatcatat cgtacagaga 360
ctgtgtgaaa gtaaagaaga tggtcagtat ggagaagttg tcagccaaat tccaaatcct 420
gatctgaacg agaacatttc tactggatta aaaccatgtg aatgcagtat ttgtggaaaa 480
gtctttgtac gtcattccct ccttaatagg catatcctag ctcactcagg atacaaacca 540
tatggagaga agcaatataa atgtgaacag tgtgggaaat tcttcgtttc tgttccaggt 600
gttagaagac acatgataat gcacagtggg aatccagctt ataaatgtac gatatgtggg 660
aaagcttttt attttctcaa ttcagttgaa agacatcaga gaactcacac aggagaaaaa 720
cctataaat gtaaacaatg tggtaaagca ttcactgttt cgggttcttg tctaatacat 780
gaacgaactc acactggaga gaaaccctac gaatgtaagg aatgtgggaa aacattcaga 840
ttttcttggt cttttaagac gcatgaaagg actcacactg gagaaagacc ctataaatgt 900
accaaatgtg ataaagcctt cagctgttcc acttcccttc gttaccatgg aagcattcat 960
actggagaga gaccctatga gtgtaaacia tgtggcaaag cctttagtcg tttgagttcc 1020
ctttgtaacc atagaagtac tcataccgga gagaaaccct atgaatgtaa acaatgtgat 1080
caagccttca gtcgcctcag ttcccttcac ctccacgaaa gaattcatac tggagaaaaa 1140
ccctatgaat gtaagaaatg cggtaaagcc tacactcggt ccagtcacct tactcgccat 1200
gaaagaagtc atgatataga ggctgggtgt agtgactcag cctataa 1247

```

&lt;210&gt; 116

&lt;211&gt; 1830

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1183192.1:2000SEP08

&lt;400&gt; 116

```

aagctttccc cgggtgtcggc ggcaggtgga atttccacgc tttatctgcg cctgcgccgc 60
gcgggattcg cggtcgagc tgaagaggtt cgcgggtccg tttcacttct ccgaaccctg 120
aggcagtggtg tgaagctggg acgccagcca tgttccagac cgcttggcgc caggaaccag 180
tgacctttga ggatgtggct gtgtacttca cccagaatga atgggcccagc ctggactctg 240
tgacagagggc cctgtacagg gaagtgatgc tggagaatta tgcaaatgtg gcttccctgg 300
cattcccatt caccacgcct gttctggtct cccagctgga gcaaggggaa ctgccatggg 360
gcctgcattcc ctgggaacct atgggacagg aggtctctag aggtatctgt ccaggggatg 420
ggccagaaac tgagaaggaa ggattaactc caaaggatca tgtgtccaaa gaaacagagt 480
ccttcagact gatggtgggg ggcctgccag ggaatgtttc ccagcacctt gactttggga 540
gcagcctaga gcagccacaa ggtcattgga taattaagac aaagtcaaag aggagacatt 600
tcacagatac ctcagccagg caccatgagg cctatgaggt caagaatgga gagaagtttg 660
agaaatttagg aaaaaatatt agcgtcagca cacaactcac tacaatcag acaaatccta 720
tggtgcagat atcttatgaa tgtggacaat gtggcagata tttcattcaa atggcagact 780
tccaccgaca tgagaaatgt cacactgggtg aaaagtcttt tgaatgcaa gaatgtggaa 840
aatacttcag atataactca ttacttattc ggcacatgat aattcacact ggaaagaaac 900
catttaaatg taaagaatgt ggaaaagggt taagttcaga cacagccttg attcagcatc 960
agagaatcca cactggagaa aagccctatg aatgtaagga gtgcggcaag gccttcagta 1020
gcagctctgt ctctctccag caccagaggt tccacactgg ggagaagctc tatgaatgta 1080
acgaatgttg gaaaactttc agttgcagct caagtttcac tgtccatcag cgaatgcaca 1140
ctggggagaa accttatgaa tgtaaagagt gtggaaaacg attaagctcc aacacagcct 1200
tgactcagca tcagcgaatt cactctgggg agaagccctt tgaatgtaag gagtgtggga 1260
aggcattcaa tcagaaaata accctgattc agcaccagcg agttcacact ggcgagaaac 1320
cttatgagtg taaagtgtgt ggtaaaacct tcagctggtg tggaagattc attctgcatc 1380
agaaactaca cactcagaag acacctgtcc aagcataggg ctatccatag ttagggccac 1440
tgtgcctctc cttttttctc tttattttca tgctttttat cagtgtcctc gctgtccttc 1500
ctgggttagac acttggtctt catcatgaac tcttctttaa gttttttgaa cctgtttccc 1560
caacatgaag tctctttatg gttagagaag accaaaaaac aaacaaacaa acttagaaac 1620
ataaaaggaa atgtaagttt ccagatttaa ggacatgct ggggtgccag caggatggat 1680
tcaagggaaa cctacattaa ggaaaattgc tgtgaattat gaaaacacca agtttgaata 1740
agaggatcat ggagaaaaaa tgagacccta aaaagtttct agagagagaa aacaaaagat 1800
cagaaatgaa agcggggcca ggcgcggtgg 1830

```

&lt;210&gt; 117

&lt;211&gt; 2274

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1183325.1:2000SEP08

&lt;400&gt; 117

```
acatttatcc atcttgtaaa atattgcttt gttttttggt gcatttcaaa cccacaggtc 60
ctgacttcag gagcaagaca cttcaagtga ggtgaggagg aggttccagg acctggatac 120
catcctttgc agggatgtaa ttcaccagag accaagattt ctagccagaa attgagggca 180
gctcaaggag agactacaga cacagcatcc tacttactgg ctatttccca gtaacaaaag 240
aaaatgatca agtcccagga atcactgacc ctggaggatg tggctgtgga gttcacttgg 300
gaggagtggc agctcctcgg ccctgctcag aaggacctgt accgagacgt gatgttggag 360
aactatagca acctcgtgtc agtgggggtat caagccagca aaccagatgc actcttcaag 420
ttggaacaag gagagccatg gacagtagaa aatgaaatcc acagccaaat ctgtccagaa 480
atcaagaaag ttgacaatca tctacagatg cactcacaaa agcaaagatg tctgaagaga 540
gtggaacaat gccataaaca taatgcattt ggaaacatca ttcacagag gaaaagtgat 600
tttcctttta ggcaaaatca tgatacattt gacttacatg ggaaaatact gaaatcaaat 660
ttaagtttag tcaaccagaa caaaagggtat gaaatcaaga attctgtggg ggtaaaggga 720
gatgggaaat ccttccttca tgccaagcat gaacaatttc ataagaaat gaacttcccc 780
gaaggtggaa attctgtgaa taaaaattca caattcatta agcatcagcg aactcaaaac 840
atagataaac cccatgtatg cactgagtggt gggaaggcct tcctcaagaa gtctcgctc 900
atctatcatc agagagttca cactggggag aaacctcatg gatgcagtat atgtgggaaa 960
gccttctcca gaaagtccgg gctcactgaa caccagagaa accacacagg agagaaacc 1020
tatgaatgca ctgaatgtga caaagcattc cgctggaaat cacagctcaa tgcacaccag 1080
aaaattcata caggagagaa gtcatatata tgcagtgatt gtggaaaagg cttcacaaag 1140
aagtctcggc tcattaatca tcagagagtt catacaggag agaaaccaca tggatgcagc 1200
ctgtgtggga aggccttctc caaaagggtcc aggtcactg aacaccagag aactcataca 1260
ggagagaaac cctatgaatg cactgaatgt gacaaagcat tccgctggaa atcacagtc 1320
aatgcacatc agaaagctca cacaggagag aagtcataata tatgccgtga ttgtggaaa 1380
ggcttcattc agaagggaaa tctcattgta catcagcgaa ttcatactgg agaaaaacc 1440
tatatatgca atgaatgtgg aaaaggcttc atccaaaagg gcaacctcct tattcatcga 1500
cgtactcaca ctggagagaa accctatgta tgcaatgaat gtgggaaagg cttcagccag 1560
aagacatggt taatatccca tcagagattt cacacaggaa agacaccctt tgtatgtact 1620
gagtgtggaa aatcctgctc acacaagtca ggtctcatta accaccagag aattcacaca 1680
ggagagaaac cctatacatg cagtgactgt gggaaagctt tcagagataa atcatgtctc 1740
aacagacatc ggagaactca tacaggggag agaccgtatg gatgctctga ttgtgggaaa 1800
gctttctccc acttgctcatg ccttggtttat cataagggaa tgctgcatgc aagagagaaa 1860
tgtgtagggt cagtcaaatt ggaaaatcct tgctcagaga gtcatagctt atcacataca 1920
cgtgatctca tacaggataa agactctggt aacatgggtga ctctgcagat gccttctgtg 1980
gcagctcaga cctcattaac taacagtgcg ttccaagcag agagcaaagt agccattgtg 2040
agccagcctg ttgccagaag ttcagtctca gcagatagta gaatttgcac agaataaaaa 2100
ccatatgaat gcagtgaatg tggtagtgct ttcagtgatc aattacatca tatgtcaca 2160
aaaacacaga ggaacaaact gatatatcca aggtggaaag cccttgaata aaaccttatg 2220
gctaataagc atatactcag agaaaaatag tatgaagtgg agactgggaa attc 2274
```

&lt;210&gt; 118

&lt;211&gt; 473

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1178269.2:2000SEP08

&lt;400&gt; 118

```
aattccccgg tcgacccacg cgaccgggta gctccgcaag gacccggacg cgaccagggg 60
aggcagaagg actctgcact tgtccaagag aagaatccag agaaatcaga gcaggtcaaa 120
tagttctaaa agccatggcc cagggactgg tgacattcag ggacgtggcc atagagttct 180
ctctggaaga atggaagtgc ctggaacctg ctcagaggga cttgtacagg gagggtgactc 240
tggagaactt tggctcactg gcctcactag gtaaggagat tgtatcacac gggccagacg 300
gaaagcaagg caacagggac tacaaccatg ggggtgcagt tccatctctg cctaaatggg 360
tgaggagagg aaactgtcac aggagacgct acaaaactaca gccatagaga gaggccccgc 420
```

cagcaatagc tgtagctact tccagagctc ctgggtcaaa tgatggagag agg 473

<210> 119

<211> 1802

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:813422.1:2000SEP08

<400> 119

gttgcttgca ttaactcaca taaatcagtg aatatgtaca cctcttttct gccattttgt 60  
gtgaaaaataa gaaattttcc tgtgtccact tgggtgaagtg tgtttttatt tcagggactg 120  
tttacattca aggatgtggc catagaattc tctcaagagg agtgggagtg cctggaccct 180  
gcccagaggg cctgttacag ggacgtgatg ttggagaact acaggaacct gctttctctc 240  
gatgaggata acatccctcc agaagatggt aaccaccctt gtacatcttt acatttttcc 300  
ttgtgagtct cttgggaggc cctgatgtgc ttgtctgaat ccgatccctg ttctaaaagg 360  
ggagattgaa actttttgat tgagagtggg aaggcttcat aatgtagcat atagacttta 420  
aacttccctt ttcttcagaa gttctgcgca taatcaggtt tgacgagtggt tggttccaga 480  
atatcagtag ttacaaaacc ttatggaaaa cgttttaaag tgtctgggat catggtttta 540  
tcctgtgtgc tttgattcag tagtttttcc agagtctgca tatttaaaat aatccctaag 600  
cattcagaag aaggcagtcg gttgactaat ttatgaaatt gttttgggaa atagttttgc 660  
agaccatgt gtaatgtcct ctcttcgtag ctgaacaaag ggctgggaat gtatccagga 720  
aaagcacagc ataccatgtt ccttttttcc tttataagtt gaaatctttc ctgagctgaa 780  
tattgtgttc attttgagc acaggaaaga gtcctagact atggaaagtg aagtgaatat 840  
agggaaaaaa tcaggtagggt cggaaatgtg gcacaaataa gcgctcagat ggccacagag 900  
taagctactc cattgagtat tgtttgggaa actctgcaaa tgggaaagat cagtgggaaa 960  
acaaaagttt atttcttatg agctctcaaa gcaagggtcca caccgctcc tactgtcccc 1020  
caacatactg ttattcaaat gtgtaaaagg ttgaactccc actctccgt gatgccccaa 1080  
ctcatacaga cacaaaagca aagggtctat catgacctac aatactgttt ggcccctctg 1140  
accatttttg tctcttggtt ttgagagggg ccgaaatggc tgtttttaag gactctcttt 1200  
tagtgtatat aatttcactc gatgtgtggt tttcagtcta tacagatcag agctggggcc 1260  
ttgtagccc tggctccaga acctattcaa ttctattcc tatttcttat acctgcaaga 1320  
acttttcagg aaaaagggaa aacagtggct ccccaaaaat gtatctgggg ttccctgaaa 1380  
tttgaagac tgtcgtctc tctacttgtt caagcccctt ggccattcct tcagttctgc 1440  
ttttgccaca tgtgtctaaa ggggaatggg caggcttttg agattgttac tcagcattaa 1500  
tgtaaagact tgacagagag atcctctttt tgattgccat tccaaaatgg gaatatatct 1560  
cattctttta ggttctcacc ttgcagcctg tggacagagc acactgcctc ttccttagaa 1620  
tcctacaaaa tccgaccctt ttattttact cccgtacctg ttatttctct cctttttcta 1680  
gctttctact ctatccggga acttagggag acctagtct tcctgattcc atattcttcg 1740  
taagatttgc ctctgaaagt tgcttttgaa atacacattc cggatgggag tgggtgcctca 1800  
cg 1802

<210> 120

<211> 1414

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1093049.6:2000SEP08

<400> 120

ggcatgcagc attatgccta tattttttta ttatttttgt agagacaggg tcttgcaatt 60  
tgtttatact tttatcccta agcaataaag ggtggttcta gatctataag ggggactttg 120  
gtattacaat tttatttaat tcataagtta ttataagtgg aagcataatt tctactgtga 180  
ttactccttc tccactata aacagtgcaa aacaaactac atcttatgta aaatacataa 240  
gacacatgga cacaccaaag agtgaattcc cagattcttg ttcttttgac aagaatcttt 300  
gggctgatgc tttgacttga ggaaaaatgt gccattacag gggttggtgt cctttgaaga 360  
tgtggctgtg cacttcacct gggaagagtg gcaggacgtg gatgatgctt agaggaccct 420  
gtacaggaac gtgatgctgg agacctacaa cagcctggta tcattacagg agttggtgtc 480  
ctttgaggag gtatgtgtgc acttcacctg ggaggagtgg caggacctgg atgacgtca 540

```

gaggaccctg tacagggacg tgatgctgga gacctacagc agcctgggtat cattggggca 600
ttgcattacc aaacctgaga tgatcttcaa gctagagcaa ggagcagagc catggatagt 660
agaagaaacc ctaaacctga gactttcagg tggaaagcaag aagcaagttt tctcaggtat 720
ttgccacagg agcctgggtg agctccagga gggttgatct ctcttgatga ctctggaact 780
gtattcccaa ttgtcaattg gacatcccta cgtatgggac ctcagatatt tcaaacatga 840
tgtgtccaag tctgtatcac ttctggccat catattgttc ttttattttt ccaaatttca 900
catcaccagt aacaaactag ctgtgatcat ggcagatagc ctggaaataa aactccccct 960
tttacccttt gcacagcaaa ttgacatcaa atcctgtttc tacttttttt tttttaacta 1020
ttgcttccct attctgtatt ctcaactgtc catctttctga tgtaggaggt catctgtttt 1080
cctcttttcc tctcctctga ctcttaagcc ctttcccatt ctctttctca ggaatggctg 1140
ttaaataatg aatattggtct tgaactttc ctgtacttag tgaacctcct tatttacacc 1200
ctgtttgtga agtggctgtg ttcaccctgg gtggacacgg aatgtttttg gcatgtacaa 1260
agagaatttt atgctgcctg tgtacagtta ttaatttgta agtacactca gctttttgta 1320
tctgtagggt taatatctgt gtatgtaagc aaacttggat gcaaaatatt tgaaataaaa 1380
tcagatgctt gcatctgtag tgaacataaa aaaa 1414

```

&lt;210&gt; 121

&lt;211&gt; 649

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:202192.4:2000SEP08

&lt;400&gt; 121

```

ctcaggtcca gcgggctcgg gcactgcttt gacagtcaca cctcccgtct cctggatttg 60
gaatcggacc ctagtctctga ccctatggga acttgatgat ttccttgatg caacaccaag 120
ccccctggaat tgaaagggac gtgcaaaggg agggacctgc cccaggggtg ttgcagtctc 180
agccgcctac atctgcagta gtaaggctct ctctgttggt acaacataca acacaaccgc 240
ccgccgttag ccactttttt tatagggtac atcataaatc acctgagatt ttagtctgtc 300
cgccggttcca ctttataact cctgcactgc ggtagatctt gatgtggaag aaatggagga 360
ctcagaacca aggatttcca agtgatttct tccaaagcac aggaatctca ctctgttaaa 420
gctgggtctgt tctaactgag atgacagtca tgtccctttc cagggacctc aaggacgact 480
ttcacagtga caccgtactc tccatcttaa atgagcagcg cattcggggc attttatgctg 540
atgtcactat cattgtggaa gataccaaat ttaaagccca tagcaatgtt ctggcagctt 600
caagcctgta ttttaaaaat atcttttgga ggcatacaat ctgtatttc 649

```

&lt;210&gt; 122

&lt;211&gt; 837

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1041854.1:2000SEP08

&lt;400&gt; 122

```

gtgggaactg cgcctcccag aagacaccgc gcgtcgtcct cctcgccctc cctgaaggcc 60
gcctcgctat cctgtgtgtg ctgctgtccg ctaagctaaa gtccactgac gtttccgcca 120
ccatgctgcc cgccgcattg ctctgtcacc cagggtctgc ccgctctggtg ctccaggcgc 180
gtacgtacgc ccaggccgcc gcctcacctg ccccgccgcg tgggcctgga cagatgtcct 240
tcaccttcgc ctccccgacg cagggtgttct ttgatgggtg caatgtccg caagtggatg 300
tgectacgct gactggagcc tttggcatcc tggcatccca tgtccccaca ctacaggtec 360
tacggcctgg gctggtagtg gttcatgcgg aagatggcac cacaactaag tattttgtga 420
gcagcggctc cgtcactgtg aatgcggact cctctgtgca attactagct gaagaagtcg 480
tgacactgga catgctggac ctccggggcag cccgggcaaa cctggagaag gcgcagtcag 540
agctgtcagg ggcagcagat gaggcagcac gggctgagat ccaaateccg atcgaggcca 600
atgaagccct ggtgaaggcc ctggagtagg cgtactttgt ctgtcaccca cagggtaaca 660
gaggcaggtc ctggggggct gaagtcgccca ccagggggca gcagtgtcct gggtactggc 720
ttaagcttc cctgggtgctg cctgccaggt catggagggc ttccccagg ttccagatgc 780
tgatcccca gatgcctctg gagagctggc cttgattggc cctcaaagca cccggac 837

```

<210> 123  
<211> 504  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1100502.1:2000SEP08

<400> 123  
cgggggtgac attcagccgc tgggttcgggg agacgaacct tccaccttag accatggcca 60  
agttcatccg taacctcgcg gacaaggcac cgctgatggg ggcggctgcc gtgacttact 120  
cgaagcctcg attggccaca ttttggcact atgctagggt tgagctgggt cccccaaccc 180  
ttgggtgaaat ccctacagct attcagagca tgaaaaatat aattcacagt gcccaaactg 240  
gtaacttcaa acacctcaca gttaaggaag ctgtgctgaa tggtttggtg gccactgagg 300  
tgtggatgtg gttttatatc ggagagatca taggcaaacg tggcattgtt ggctatgatg 360  
tttgaagacc aatctttaac ttctgatttg agttcttctc ttggaccatg tgtaatgaga 420  
ctgctatctg aataaaatgc tagctgtgca gtggcagtgt ttctctctgc aagtactgca 480  
cagtgcacgt atttcctgat aggt 504

<210> 124  
<211> 925  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:726414.1:2000SEP08

<400> 124  
caggcagggtc ggcgaggagg gcaggaagac cgaagaccga gccgtgctgg ggctcgggac 60  
gggaacggga caggacccca aaatctcaga tccttcgggc ccgcccgcgc gtgcccgtcg 120  
acgcgtcggt cttgccggcc gcgcctcacc tccgccctct cctcctccag ggggatcgga 180  
tacgccacag gctgcgcgat ggtgctgtgg gtcttcggct acggctccct catctggaac 240  
cccggcttcg acttcgacga caaaatcctc ggcttcatca agggctacaa gcgcaccttt 300  
aatctcgctt gcattgacca cagaggcaca ccggagcatc cggcgaggac ctgcacgctt 360  
gaaaccgacg acgaggccat atgctgggga attgcatatt gtgtcaaggg tgggtccagaa 420  
aaagagctaa aagcaatgca gtacttgagg agaagagagt gtgagtacga ccagaagata 480  
tcatttgatt tctacaagga aggagatccc ttgaaaccag ctgtgacagg cgtcttagtg 540  
ttcgtttcca ctccagatcc aattggcaac aagtactatc ttggccctgc tcctttgcag 600  
gatatggcaa ggcaaattgc tacagccaat ggccctactg gctataatag ggattacctg 660  
ttctcaatgg agaaggcatt agccagcatt agccacgaag atgattcgat catagagctt 720  
gcaaacgagg tgaggaaggt gctcaacaga acaaaggaga ccaagatcac tgggtgccaat 780  
gttcccctga aatcacatgc cctcttgtg cactgtccg cgcttcaga aggcaccgtt 840  
gtggactcga gatagcgaga gccttggacg ttggacctct actcggaatc tacttccaag 900  
ctcttgcattg tttgtggcca tcatg 925

<210> 125  
<211> 3194  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:400517.4:2000SEP08

<400> 125  
gggggctggg agctggctgt cccaggagca ggatgccctg gccagcccc agccctgggt 60  
gaagaccag ctgttcatgt gggagcctgt gctgtttggg acctgggatg gtgtgttcac 120  
atcctgcatt atcaacatct ttggggttgt gctcttctcg aggactggct ggctgggtgg 180  
aaacacagga gcgctcctgg gcatgttctt ggtgtccttc gtcatectgg tggccctcgt 240  
cacggtgctg tctgacattg gcgtcgggga gcgcagcagc atcggcagcg gtggcgtcta 300  
ctccatgatc tcctcggtcc tgggtgggca gacgggaggc accatcgggc tgctctatgt 360

```

gtttggacag gtgccatgta tatcacccggc tttgtctgaat ccatctcgga tttgtctgggc 420
ctcgggaata tctgggctgt gcgaggaatt tcagttgcgg tgcttctggc cttgcttggg 480
cattaacctc gcaggtgtca aatggataat ccgcctccag ctgctgttgc tgttctgtct 540
ggccgtgtcc aactggact ttgtgggtggg ttctttcacc cacctggacc cagaacatgg 600
tttcattgga tattcaccgc aactgctaca gaacaacacg ctgcccgcatt acagcccggg 660
ggaatctttt ttactgtct ttggggtttt ctcccagcg gctacaggag tcatggccgg 720
cttcaacatg gggggctgcac ctcagggagc ctgccgacca gcattcccct gggctccctg 780
gcagctgttg tgcattctga agaagcttct ctgcccttca ttgggtgggt tctgtacatc 840
tgtcttctgc ttctctcttg ggcgccatct gcactcgaga ggcccttcgc tatgacttcc 900
tgcatagcgg acaacaggta tccctcatgg gcttctgtt ccttttgggc ttatacatct 960
cgtccctggc ttcttgcctg ggaggacttt atggagctcc cctgtatcct gctgtgcatt 1020
gccaggaga aagtgatccc tgcacttgcc tgtctgttga cacagggaag gggccaaaca 1080
aaacaccccg tggcttgcca tcttgctga ccagcttggg gacctggcc gttcgttttt 1140
gtgggtcaag tgaacgttct ggccccatc acttcatgct gacatacgtt 1200
gcagtggact actcttactt ctccctgtcc atgtgttctt gcagcctgac cccggtgcct 1260
gaaccggtgc tcaggggagg gcgcagaagg cctccactgc tctgagcacc tgccttaga 1320
gaaagctccc agttacggct ctgagggacc tgcccaaaga gtcttggcag ggcacgctac 1380
tggaattcca ccaaggacat ggatcagctc ctccagctaa ccaggaagct tgagagtacg 1440
cagcccaggc aaggagaggg taacaggacc ccgagaaagt cagaagagga aaagcaagaa 1500
ggcgaccaa gcagacccta caagatagct tctcttggg cctcaaatac ccctccttct 1560
ttccctgtcg agatctctga ccaggttgcc cgctgcctcc tgggaggggc aggagtcctg 1620
ctggaacaag cagacttcca agagcgaagg gactcagcct gagggaacat atggagagca 1680
acttgttctt tgagctgggc aaccctaatc agagtccag tggagaagat ttcttccatg 1740
aagtctcagg ctccaagaac agagatgtct ggagaagatc cacttcttcc tataccaca 1800
tgtgcgacc cggggtctcc cgtgttggcg ggctgtagag gtcccgttct catcatgttc 1860
tgtgatacag gtggagtgtg taccctgggt aacatgggtg ttgctgcat cgtgtatttc 1920
tacattggcc gggccagtc agggcttcac ctgtgatcag cctccaactt cagctttttc 1980
cggtggtatg ggtctctctt gctccctccc tgcaggagct tgcggtcccc tcaggagcag 2040
atcatcttgg gcgcgctccc tggctaagg tgcattggag atgactcagc tcaccaggga 2100
gaatgcagac ttgcgcactc gggatcgcta ccaccactcc tccctcgtga accgggagca 2160
gctgatgcct cactactaga tgcagtgtcg ggaccttctt cttttggagc tgtcccatgt 2220
acagtggacc caagcccagg accttctgtg agctgcttct ccaacctgag aaactcaaga 2280
cccatcctcc cgctagtcac tttggacaat ggaaatctac attttctttt cccttttttt 2340
tttttttggg agacagagtc tgccttgtgt aaccagggtt ggagttccag gggcacaatt 2400
cttgggtctc cactgtgcaa cctctccaac ctctcccga gtttacaagc aatttcttcc 2460
ctgtgactct cagcctctcc tgtgagtata ggcgtgggat tatagcctca tgtgccacgc 2520
accccagat atttttggta ttttatacgg gagacggggt ttccacttag tggtagggcca 2580
ggctgtgggt ctcgaaactc tgtgacctcg ttgattccac ctgggtgttct acagccttcc 2640
caaaggtacg gggaattaca ggcgttgag ccaccatgtg ctgtggccag aaatctctag 2700
tgtttcctta gaacttgggt gaagagggaa aaggaccaa aagggaagtc gggatttctg 2760
aggccactt tctcaccggg ggttgactct caggactggg gtgctagccc ctctgcaaac 2820
acgacaccca gtctgaaaga gctctaatac aggtactaag acttagcggg gggaacacaa 2880
gatctctgca caaccagat tgcctgtgca gcgggccacc aataagtgtg tgttttatat 2940
tttattgtat ctattttatt aacaaataaa taatacactc acatgtttcc acacccaaaa 3000
gatatataga gtgaaaagtc tccctccaat ccttattctc cactgtctca gttcctgtct 3060
accctggcct agaggtaacc actattctca tccgtacatc cagcatttct ttatggatat 3120
acaagccaat gggatgtag agtctcacat tctctcttta tacacaaaag ttagcacggt 3180
atacacatgt tctg 3194

```

&lt;210&gt; 126

&lt;211&gt; 1110

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1078917.1:2000SEP08

&lt;400&gt; 126

```

tgacaagatt gtgaatggag aagcagatgc catgagcttg gatggaggtc atgcctacat 60
agcaggccag tgtggactag tgcccgtcat ggcagagaa tatgatattc cttcgtgtac 120
aaacccacaa tcagatgtct ttccctaaagg gtattatgcc gtggctgtgg tgaaggcatc 180
agactccagc atcaactgga acaacctgaa aggcagaag tccctgccata ctggagtaga 240

```

```

cagaaccgcc ggctggaaca tccctatggg cctgctgttc agcaggatca accactgcaa 300
gttcgatgaa tttttcagtc aaggctgtgc tcctggctat aagaagaatt ccacctctg 360
tgacctgtgt attggcccag caaaatgtgc tccgaacaac agagagggat ataatggtta 420
tacacggggc tttccagtcg ctcgttgaga agggagacgt agcctttgtg aagcaccaga 480
ctgtcctgga aaacacgaac ggaaagaaca ctgctgcatg ggctaaggat ctgaagcagg 540
aagacttcca gctgctgtgc cctgatgggt acccaagaag cctgtaaccc gagtttcgcc 600
actggcaact gggccaaggg ttcaaaacat ggtgtggtct tcacgaaaaa gagaaggcag 660
cccgggttag cactgtgctg actgcccaga aggatttatt ttggaaagggt gacaaggact 720
gcactggcaa tttctgtttg ttccgggtctt ccaccaagga ccttctgtta cagagatgac 780
accaagtgtt tgactaaact tccagaagggt accacatatg aagagtactt aggagcagag 840
tacttgcaag ctgttggaag cataaggaag tgttcaacct cacgactcct agaagcctgc 900
actttccaca aaagttaaaa tccaagagggt gggtgccact gtgggtggagg aggatgcccc 960
ccgtgatcca tgggcttctc ctgggctcca tgccctgagc ggctggggct aactgtgtcc 1020
cgtttcactg ctgtgtgtta ccacatacac agaggcacaa agataaaaaa tgactgttga 1080
ctttaaaaaa aaaaaaaaaa cggccgccga 1110

```

&lt;210&gt; 127

&lt;211&gt; 1147

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1012560.1:2000SEP08

&lt;400&gt; 127

```

tagtggatcc ccgggacgca ggaattcggc acgagaataa attgaagatg gacatgctgt 60
aactttctaac acatgggttaa gcttctcaga ttacctaata ctaagaaagg tttgctttta 120
gaaatagtgc tcccttcaga atggaagaat ttatctgcct ctattttgat gtggatcaga 180
gctaagatgg ctgactaaat aaacatgggg gactggaatc tccttggcag atactctgga 240
ggaagtccac atccactcca ccatgattgg aaagatctgg ctcaccatcc tgttcatatt 300
tcgaatgctt gttctgggtg tagcagctga agatgtctgg aatgatgagc agtctggctt 360
catctgcaat acagaacaac caggctgcag aaatgtatgc tacagaccca ggccttatcc 420
tatcatccca tcattagata ctgggttctg cagggtgatat ttgtgtcttc accatccctg 480
gtctacatgg gccatgcatt gtaccgactg agagtctctg aggcaagaga ggcaacagga 540
tgaaagctca gttaagagta gaactggcgg aggtagacgt ttgaccatgc ctagggatcc 600
ggcaggcgat tggagccaag agctttgtca gctggagaaa aggaaactaa ataaagctcc 660
actcagagga accttgcttt gcacttaatg tgatacacat tttcactcgc tctgtgggtg 720
aagttggatt catgattgga cagtaccttt tatatggatt tcacttagag ccgctattta 780
agtgccatgg ccaccgtgt ccaaatataa tcgactgttt tgtctcaaga ccaacagaaa 840
agacaatatt cctattattt atgcaatcta tagccactat ttcacttttc ttaaaccattc 900
ttgaaatttt ccacctaggt tttaaaaaga ttaaagagg gctttgggga aaatacaagt 960
tgaaagagga acataatgaa ttccatgcaa acaaggcaaa acaaaatgta gccaaatacc 1020
agagcacatc tgcaaattca ctgaagcgac tcccttctgc ccctgattat aatctgttag 1080
tggaaaagca aacacacact gcagtgtacc ctagtttaaa ttcacacctt gtattccagc 1140
caaatcc 1147

```

&lt;210&gt; 128

&lt;211&gt; 2440

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:427997.4:2000SEP08

&lt;400&gt; 128

```

cgcgtacagg aggcggcggc ggctcccagt caccggcccc cgacggcgag cgcacgatgc 60
actgcctggg cgccgagtag ctggtttctg ccgacggcgc ccctaggcaa agggagtggc 120
gaccccagat ttataggaaa tgcacagata cgcatgggtt attcctgttc tttctctttt 180
ggactgggtt ggtgtttatc atgggctact cgggtgtggc tggagccgcg ggaagactcc 240
tctttggcta tgacagcttt ggcaacatgt gtggcaagaa gaactcccc gtggaagggg 300
cccctctttc agggcaggac atgaccctaa aaaaacacgt gttctttatg aattcctgca 360

```



```

acctggaagt caaagggtacg cagctcaacc gcatggccct ctgtgtatcc aactgccctg 420
aagagcagct tgactccctg gaagagggtcc agttctttgc aaacaccagt gggtccttcc 480
tgtgtgggta tagtttgaat tcttccaact ataccacag tccaaaagca gactcactgt 540
gtcccaggct accagttcct ccaagcaagt catttccctt atttaaccga tgtgtccctc 600
aaacacctga gtgctactcc ctatttgcac ctgttttgat aaatgatgtt gacaccctcc 660
accgaattct aagtgggaatc atgtcgggaa gagatacaat ccttggcctg tgtatcctcg 720
cattagcctt gtctttggcc atgatgttga ccgttcagat tcatcaccac tccttctggt 780
tcacattttt catttacatt gtgttatttt gggattgttg tttagtctgc ggtgttttat 840
gcgtagtgtc agtattatga ctataccaac gaccttcagc gatagaattg gacacagaac 900
agggaaaata tgaagtgcgt gctgggggtt ggctatctgt atccacaggg catcacggct 960
cagtgtctgt cgtcttgata ttttgttctc agaaagagaa taaaattgga cagtttgagc 1020
ttttccagaa tcacaaataa agccaatcag cagttggctc ccttccctggg ctgggttccaa 1080
gccaaactggg tgggacgatt tggccatccc tcattgttct tcgtgggtcc tctgggtggc 1140
tgtgcgtgct gaggcctggg aactgcagga gcgtggccag gttatgggaa ggccggcgaag 1200
tggagagtat aagccccttt cgggcattcc ggtaacatgt gggtcgtacc atttaattgg 1260
cctcatctgg actagtgaat tcacctctgc gtgccagcaa atgactatat gctggggcag 1320
tggttacttg ttatttcaac agaagtaaaa aatgatcctc cgtgatcatc ccagccgttt 1380
cgtctcgtcc ccattctctt cttctaccat ccaaggaacc attgtgaaag ggtcattttt 1440
aatctctgtg gtgaggattc cgagaatcat tgcattgtac atggcaacaa cgcagctgac 1500
aagatacagc atcgggtgcat tcgtaccagg tacctgttcc ggatgactgc ttacggctgt 1560
tcctgggtgtc tggacaaaat acctcgctcc atctcaagcc agaatgcata tactacaatc 1620
ggcctattac atgggacaga tttcgtgtac atcagcaaaa gatgcattca aaatcttgtc 1680
caagaactca agtcacttta catctattaa tctgcttttg agacttcata atttttctag 1740
gaaagggtgt agtgcgtgtt ttcactgttt tttggaggac tcatggcttt taactacaat 1800
cgggcatttc cagggtgtgg gcagtcctc tgttattggg agcttttttt ggccacttta 1860
gttagcccat agttttttta tctgtgtttg aaactgtgct ggatgcactt ttctgtgtt 1920
ttgctgttga tctggaaaca aatgatggat cgtcagaaaa gccctacttt atggatcaag 1980
aattttctgag tttcgtaaaa aggagcaaca aattaaacaa tgcaagggcc acagcaggga 2040
caagcactcc attaaaggaat gaggagggga acagaactcc caggccattg tgcagataga 2100
taccacttta ggtatctgta cctgggaaaa catttccctg ctaagaggcc atttacagaa 2160
tagaagatga gtatccacta gagataagtt tgtgaatttt tttttctcaa aagacctaat 2220
aaaaccctat tcatcactca ttgtactttg tcattattgt tcgaaccagg ataacaatac 2280
tggaactata ttagtttacc tttttttgta caaattagga cagaaaaact cttctaaaac 2340
catgtttata tgcatcaact taaaaagta actatgtaag aactgaggta agtttgtaag 2400
tgcaacaata ataaataaac ctttttaaga taaggaaaaa 2440

```

&lt;210&gt; 129

&lt;211&gt; 2225

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:197899.1:2000SEP08

&lt;400&gt; 129

```

gaagcctgag cggccgaact cggcagctcc aacccaacte ggcttaacte cgcctcaccg 60
agcccagctc aagactctgt gctccctagg tttgcaacag ctctctgac atcttcttca 120
attcctgcta ggatgccgtg gcaagcattt cgcagatttg gtcaaaagct ggtacgcaga 180
cgtacactgg agtcaggcat ggctgagact cgcccttgcca gatgcctaag caccctggat 240
ttagtggccc tgggtgtggg cagcacattg ggtgcaggcg tgtatgtcct agctggcgag 300
gtggccaaag ataaagcagg gccatccatt gtgatctgct ttttgggtggc tgccctgtct 360
tctgtgttgg ctgggctgtg ctatgcggag tttgggtgcc gggttccccg ttctgggttcg 420
gcatatctct acagctatgt cactgtgggt gaactctggg ccttcaccac tggctggaac 480
ctcatcctct cctatgtcat tggtagagcc agtgtggccc gggcctggag ctctgctttt 540
gacaacctga ttgggaacca catctctaag actctgcagg ggtccattgc actgcacgtg 600
ccccatgtcc ttgcagaata tccagatttc tttgcttttg gcctcgtgtt gctgctcact 660
ggattgttgg ctctcggggc tagtgagtcg gccctgggta ccaaagtgtt cacaggcgtg 720
aaccttttgc ttcttggggt cgctcatgat tctggcttcg ttaaggggga cgtgcacaac 780
tggaaactca cagaagagga ctacgaattg gccatggctg aactcaatga cacctatagc 840
ttgggtcctc tgggctctgt aggaatttgg gccttttggc ttcgagggaa ttctccgtgg 900
agcagcgacc tgtttctatg catttgttgg tttcgactgt attgctacca ctggagaaga 960
agcccagaat ccccgagcgt ccatcccgat gggcattgtg atctcactgt ctgtctgctt 1020

```

```

tttggcgat tttgctgtct cttctgcact caccctgatg atgccttact accagcttca 1080
gcctgagagc cctttgcctg aggcatttct ctacattgga tgggctcctg cccgctatgt 1140
tgtggctgtt ggctccctct gtgctcttct taccagcctc ctgggctcca tgttcccat 1200
gcctcgggtg atctacgcga tggcagagga tggcctcctg ttccgtgtac ttgctcggat 1260
ccacaccggc acacgcaccc caatcatagc caccgtgggtc tctggcatta ttgcagcatt 1320
catggcattc ctcttcaaac tcaactgatct tgtggacctc atgtcaattg ggaccctgct 1380
tgcttactcc ctgggtgtcga tttgtgttct catcctcagg tatcaacctg atcaggagac 1440
aaagactggg gaagaagtgg agttgcagga ggaggcaata actactgaat cagagaagtt 1500
gaccctatgg ggactatttt tcccaactcaa ctccatcccc actccactct ctggccaaat 1560
tgtctatgtt tgttcctcat tgcttgcctg cctgctgact gctctttgcc tgggtgctggc 1620
ccagtgggtc gttccattgc tttctggaga cctgctgtgg actgcagtggt ttgtgctgct 1680
cctgctgctc attattggga tcatttgtgt catctggaga cagccacaga gctccactcc 1740
ccttcacttt aagggtgcctg ctttgcctct cctccacta atgagcatct ttgtgaatat 1800
ttaccttatg atgcagatga cagctgggtac ctgggcccga tttggggtct ggatgctgat 1860
tggctttgct atctacttgc gctatgggat ccagcacagc ctggaagaga ttaagagtaa 1920
ccaacctca cgcaagtcta gagccaaaac tgtagacctt gatcccgga ctctctatgt 1980
ccactcagtt tgacatcgtc acacctaaat gctgtctggt ccctgcaca ataatggaga 2040
gtactcctga cccagtgac agctagccct ccctgtgat ggtgggtggt gatactaata 2100
cagttctgta cgatgtgaag gatgtgtctt tgcattttct tgcattttt aaccogtctg 2160
cttctaaatg atgtctagct gcttaccac tttaaaaaat gatattaaaa gaaagtagaa 2220
aaata 2225

```

&lt;210&gt; 130

&lt;211&gt; 1218

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:334199.1:2000SEP08

&lt;400&gt; 130

```

ggaaacaaac tctgttttag tgagtcaatg tatatttctc acataaagaa tttgatctca 60
gaaatgctg tttactcaca tactttcaag attttaaatt tgcttttgat aaataaatgc 120
tgcacagtag gccgtaacca ttgacaatat tcttggaatg ttcccaaaat cctgttagat 180
acgaaagtca gtatctgaag gaatgattcc actctgaaaa acctcctgag ccaccacaga 240
agcaaaaggg tgtttagctg ctgaaacata agcttgaacc aaccaaggat tttcaggacc 300
tgtttggaat acaagttctt tccctcctac acctgctgcc tctagggttaa tgaatgcacg 360
aatcaagcta gcccaggggt gctgagtaat gaaacatga ctggcttgca agacattttc 420
ctcagcacca ttaaagagaa atatgacagc atgatgcaag gttctgaaag atgttgacaa 480
gacgcgaagg acttccagca tcaactgagca gctaaactgca tcatcactgg cacctgggtga 540
gtttgctact gagtcaaaat gacaattagc caagacagca tgctgggctc catctctggg 600
ttccagcttt accacaacat tgggtgatgtt gtcataatag cttgtaaaaac ctcccaagaa 660
atcaatgcta aaagagcctg tgggcccgtt tacatctact gaaatcttat gaaggctgtt 720
gctttgcact tcaatcagtt taatctgttc caaaaggtag tgcacgggtc gaatttcatt 780
ttctggactt cctgtagtcc tggggccaat ggagggttat tgttcaagat aatccctggc 840
ttggagcgcg tcgaactccc cgcggtgtcc agcggccccc cgtagcacga gctgctgcag 900
cgagagctgc accagcgtcc gcagcgcgat caggtagagc gcgagcccca gcgcggcgcg 960
cacctcagac agcccgggtc ccgcgcccct gctcgcgcgc ccgctacccc cggggctcct 1020
cttccgcgtc ctccgcgcgc cgctgcaccc atccaccaga ggctcctgcg ctccggcctc 1080
cctctccggc ggtggcgcgg ccgcgcgtcc ctctcgacgc tctactccga cgcggtgcgc 1140
cctcacagca gccgactcag accccactcc atggccacga gcctcagctg ccagcccaac 1200
cgccccaacc cgcgacag 1218

```

&lt;210&gt; 131

&lt;211&gt; 606

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:334345.1:2000SEP08

&lt;400&gt; 131

```
gcgtccggga gccgtgcaca acaccaagcg catggactct gacccagtg cagtgtctgt 60
aggggaagagg agccatgggg cttcgggcag gcccacacct gcttctgctg ctgtggctgc 120
tgccagaatg cggccactcc aaggaggcgg ggaggattgt gggaggccaa gacacccagg 180
aaggacgctg gccgtggcag gttggcctgt ggttgacctc agtggggcat gtatgtgggg 240
gctccctcat ccacccaacgc tgggtgctca cagccgcccc ctgcttcctg agggtgactc 300
cggggggcgg ctggtctgcc ccatcaatga tacgtggatc caggccggca ttgtgagctg 360
gggattcggc tgtgcccggc ctttcgggc ttggtgtctac acccaggtgc taagctacac 420
agactggatt cagagaacct tggctgaatc tcaactcagg atgtctgggg cccggccagg 480
tgccccagga tcccactcag gcacctccag atcccaccca gtgctgctgc ttgagctgtt 540
gaccgtatgc ttgcttgggt ccctgtgaac catgagccat ggagtccggg atcccccttc 600
tggttag 606
```

&lt;210&gt; 132

&lt;211&gt; 1286

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:228092.1:2000SEP08

&lt;400&gt; 132

```
ccgccctgca ggcggttgcg ttccccgtcg ttaccctctt tctcttcccg acgcgtgagt 60
tagggccgtaa tgccttgggt gctctcagcc cccaagctgg ttcccgtgtg agcaaacgtc 120
cgcgccctct caggagcttc ctggacgatt taacaggctc tcaagcgagt cttcaaagaa 180
gtgaaaagaa aatggatccc tttcatgaaa aatttcttta aaaacatcat ctggcttatg 240
gaatgtgaag ccgtactcac attcatcatc aaaatgactt ccacctccgt ttaatccttc 300
tgtgccaat ttatcataaa tgtcccgttt ctcacatatt gataatacct cgtatgcctc 360
agctactttc ttgaattttc tctctgcttc ttctttattt tctggatttt tatcaggggtg 420
ccatttaagt gccactttat gataagcttt tttaatgtcc tcaggtgaag catatctttg 480
cagtcctaga acttcatagt aatccacat gttttaacag atagttggaa gtgggtgtgc 540
tgggtattga gaaccgtggt ttcctcagct caggctctgc tgggtgggtg gatagaggta 600
gtgactgcaa ataggtacac ttttatgtta aagacaatgt tattaccatg tgctaagaga 660
tgaggatact gattatgtgg caagcagagt tacaatatg tcagagaggg ctgcattatg 720
gtgtcattct ggggtcagtg tagggactgg accttccaaa tagagctagt ccagagtcgt 780
accagaccat cctgaacct ctggactgtt tcccctgtat gtttccctgg aagcttcagg 840
cagtgcctca taagccaatg gaatctgttg ctaatagcca cagcatatcc cttgcataat 900
atgacctcta gattactgcg ccttaattgc tccccagctc ttctatgctt tggtttagaa 960
aaatgaagta ctgacttacg ggtgaagaaa gtattcaaac agttgacata ttatttcag 1020
tcaagaaaca gttcagaggg agatacaaac aagtaactta gttacaatat aatagttatg 1080
atgagaggaa gtactggatg ctaaacaatt atatgagaga cagctcaggc tgggggtgtc 1140
aatgaaagcc tcttggagga agtagcctga tatgttaact ttctgcatgc cagtggtaag 1200
tcatttgaga ttatagtatc tttgtccacc caaggtttac tgctgtaaaa catgtaatta 1260
aagattggaa agttgttgggt tatgta 1286
```

&lt;210&gt; 133

&lt;211&gt; 801

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:098580.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 49

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 133

```
gttataaaag agaattcac agctcagatg gtgctgggaa ttgtttggnt caccactgac 60
agtggaaaaa gccagtgcag atactcatgg tcggctcttg caaggtaaca tctgtaata 120
```

```

tgctgttaca aaggctcatg tgggaaaagg acttcattgc tttcaaactc tccacaccac 180
ataatgtttc ttggagacat gaaacaaatg gctccgtcct catttcccaa attatctact 240
acttcagaga gtattcttgg agtcatcatc tagaggaaat ttttcaaaag gttcaacatt 300
catttgagac cccaaatata ctgacccagc tgcccacat tgaagacta tccatgacac 360
gatatttcta tctctttcct gggaattaaa actcataaga agcaactcag gaaaatgaag 420
gcattatggt gtctatgaat ctgctattag aaaagggtcaa tttttgatta agaataattt 480
tataccaaac tgctcagctg tccacctttt caaataaccag gctaaaatca atctcatatt 540
tttgctaaat taatgcttaa taaattttta ttataaaatt gtgcttggtt tctaattttg 600
tcataaaatta agtcatgctt ataggagaaa aaatttaagta ttaaaaatca taaaaaatga 660
agaaataatc atcatcaacc tttttaaaac tattcataaa aactcaagtt ttttatatag 720
cttatttatt ttgtatatcc aatttcgtgt gcttttatta tattgttata acattaactac 780
tagtaataaa actatcctaa a 801

```

&lt;210&gt; 134

&lt;211&gt; 709

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:969572.1:2000SEP08

&lt;400&gt; 134

```

ctggcatgaa ttataaaatt atctactcca ttgtgcaaac aaattgttca aaggaggatt 60
ttccttttct cgtgaagac tgtgtacccc ttccctatgg cgatcatggt gagtgtaggg 120
gtcataccta cgtggatatt cataacacaa ttgctggcct ctacacagagc tgtgaccttt 180
atccaggaga cgatttggtt tctactacttc ccaagaaatg ctttggctgc cccaagaaca 240
tacctgtaga cagcccggag ctcaaggagg cacttggtca ttcaattgca cagcttaatg 300
cacagcataa tcactttttc tatttcaaga ttgacaccgt gaaaaaggca acatcacagg 360
tggttgctgg aacaaaatat gtgattgagt tcatagccag agaaaccaat tgttccaaac 420
aaactaacac agaactgaca gcggactgtg agaccaaaca cctcgggtcaa agcctcaact 480
gcaatgctaa cgtgtacatg agaccttggg agaacaaagt cgtcccgact gtcagatgcc 540
aagcgctaga catgatgatt tctaggcctc caggattttc acctttccgg ctggtgcaag 600
tacaagaaac taaagaagga acaactaggc tcctaaactc atgtgagtag aagggcagac 660
tctcaaaggc aggggcaggc ccagcgctg accatcagggc agaagcttc 709

```

&lt;210&gt; 135

&lt;211&gt; 1759

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:196958.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 63, 76, 137

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 135

```

tgggggttag gctccagggc ttctgtcgag accaaggatg cccaaatata tgggtggtgg 60
gtncctgcata cctggncctt gggcagaacg aagggtatac agcctgggccc accaggataa 120
gtccagaacc caccagnagc tgaggacaga cagaaggacc acggagggggg tgacgggctg 180
gtgtgaggat tgggtccctt gggccaggac tctcctctct tctccctgct ggctccagac 240
cagagtccaa gccctaggca gtgccacct taccagccc agccttgaag acagaatgag 300
aggggttttc tgtctccagg tcctgtcctt tctggtgctg ggagctgctg ggactcaggg 360
aaggaagtct gcagcctgcg ggcagccccc catgtccagt cggatcggtg ggggcccggga 420
tggccgggac ggagagtggc cgtggcaggg gagcatccag catcggtggg cacacgtgtg 480
cggggggctg ctcatcgccc cccagtgggt gctgacagcg gcgcactgct tccccagtg 540
cctcccaga gtggcgaccg ctacaaggag taaggggtgcc gctgctggac tcgcgcacct 600
gcgacggcct ctaccacgtg ggcggcgacg tgccccaggg tgagcgcat gtgctgcctg 660
ggagtctgtg tgccggctac ccccagggcc acaaggacgc ctgccagggtg tgcaccagc 720

```

```

ctccccagcc tccggagtc cctccctgtg cccagcacc tccctccctg aactccagga 780
cccaggacat cccaactcag gctcaggatc ctggcctcca acctagaggc accacgccag 840
gggtctggaa cctgagaac tgaagtccct ggagggtctg gacctaggct cctctttctc 900
ctgcagggtg attctggggg acctctgacc tgcctgcagt ctgggagctg ggtcctgggtg 960
ggcgtggtga gctggggcaa gggttgtgcc ctgcccaacc gtccaggggt ctacaccagt 1020
gtggccacat atagcccctg gattcaggct cgcgtcactt ctaatgctag ccggtgaggc 1080
tgacctggag ccagctgctg gggtcctca gctcctggt tcatccaggc acctgcctat 1140
acccacatc ccttctgcct cgaggccaag atgcctaaaa aagctaaagg ccacccacc 1200
cccccccac cactcctgc ctctctcct ctttgggat caccagctct gactccacca 1260
accctcatcc aggaatctgc catgagtcac agggagtcac actccccact cccttctctg 1320
cttgtattta cttttcttgg ccctggccag ggctgggctg aaggcacgca gtgatgggca 1380
aaccaattgc tggccatctg gctgtgtg ccatcttttt ctggagaaag tcagattcac 1440
agcatgacag agatttgaca ccaggagat cctccatagc tggctttgag gacacgggga 1500
ccacagccat gagcggcctc taagagctga gagacagccg gcagggaatc ggaacctca 1560
gacccacagc cgcaaggcac tggattctgg cagcacctg aaggagctgg gaagtaagtt 1620
cttccccagc ctccagataa gagccccgcc ggccaatccc ttcatttcaa cctaaagaga 1680
ccctaagcag agaacctagc tgagccactc ctgacctaca aagttgtgac ttaataaatg 1740
tgtgttttaa gctgcaaaa 1759

```

&lt;210&gt; 136

&lt;211&gt; 1033

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1087811.1:2000SEP08

&lt;400&gt; 136

```

gaaagaaaga aatgctaggg gaaaaatgttt taactagtca ttcttcccag tagctaata 60
agctgacttt taaaaagaag gctgtgagct ttgcagatgc tgctgccgcc cagggccccc 120
tgcttccagc catggtcaac cccaccatgt ttttccacat tgctgtcgat ggcgagccct 180
tgggctgtgt ctcccttcgag gtaaggggcc tggaaagcaa gaagtgaactg ctcatcta 240
ccataaaact atgttaacag attggagctg tttgcagaca aggttccaaa gacagcagaa 300
aatttccatg ctctgagcac tggagaaaaa ggatttgggt ataagggttc ctgctttcac 360
agaattattc caggggtttac gtgtcagagt ggtgacttca cagccatac agcattgggtg 420
gcaagtccat ctgcaggag aaatttgatg acaagaactt catcctgaag catacgggtc 480
ctggcatctt gtccatggca aatgctggac ccagcgtgaa cgtttcccag ttttttatct 540
gccttgccaa gatgccaaga cagagtgggt ggattgcaag catgtgggtct ttggcaaggt 600
gaaagatggc atgaatattg tggaggtcat ggagcacttg ggttccaaga atggcaagat 660
cagcaatcag caagaagatc accattgctg actggacaac tgcaataaat ttgacgggtg 720
tttctcttaa aacaaaaaaa aaaatactgt gacagaccaa ggtaaattgt ttttgataat 780
tacagaagtg tgatttctta atcagcagct gtcaaacata gtgttcttta atttaaaagc 840
ttgaacacta aattataaat tggagagggt ggaataatta cggtcataatc tctagaaaca 900
caagtcttta gtagcaaaaa agaaattagc aagagaagaa aactgttcag tactttgaaa 960
ggaaaaagtt ttcagtgata gtttttttag atgaaaatta acatgataaa gaagtgcag 1020
gccaggtgtg gtg 1033

```

&lt;210&gt; 137

&lt;211&gt; 554

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1327885.1:2000SEP08

&lt;400&gt; 137

```

gggccccag agaccccgcc atggttcgtg cgggcccgt ggggacgcat ctccccacgt 60
ccagcctgga catctttggg gacttgagga agatgaacaa acgacagctt tactaccagg 120
tcctaaactt tgccatgatt gtgtcttctg cgtcatgat atggaaaggc ttaattgtgc 180
tcacgggcag tgagagtccc atcgtgggtg tactgagtgg cagtatggag ccggcctttc 240
acaggggaga tctgctgttc ctcacaaatt tccgggagga tcccatcaga gctgggtgaaa 300

```

```

tagttgtttt taaagttgaa ggaagagaca ttccgatagt tcacagagta atcaagggttc 360
atgaaaaaga taatggagac atcaagtttc tgactaaagg agataataat gaagttgatg 420
atcgagggtt gtacaaagaa ggccagaact ggctggagaa gaaggacgtg gtaggaagag 480
ctcgagggtt cttaccatat gttggcatgg tcaccattat aatgaacgac taccctaagt 540
tcaagtacgc tctt 554

```

```

<210> 138
<211> 793
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:449393.1:2000SEP08

```

```

<400> 138
cagccatgtc gctccgcgtg ctgaacccta acgccgaggt gctcaacaag tcggcagcgc 60
tgcacatgaa catcaacgcc gccaaagggc tccaggacgt gctcaagacc aacctcggcc 120
ccaagggcac catcaagatg cttgtgggtg gagctgggtg cctgaagctg acgaaggatg 180
ggaacaccct cttgaaggaa atgcaaattc agaaccacac tgcaatcatg attgcgagga 240
cggctgtggc acaggatgac acaagtgggtg atggcacaac ctccactgtg ctatttatcg 300
gggagctaat gaagcagtcct gagcgggtgca tcgatgaagg aactcatcca agattcttgg 360
ttgatggttt tgatgttgct aagagagcgt gtcttgattt tctcgataag ttcaagacgc 420
cagtcgttac tgggtgaagaa cctgacagag ataccttaaa aatgggtagca agaacaacac 480
ttaggacaaa gttgtatgaa ggggttggtg atcagttgac tgatattggt gtgaatgcag 540
ttctgtgcat tcgcaaacca gatgaaccaa ttgatctatt tatgggtggaa ataatgcaca 600
tgcgccacaa atttgatggt gacacccgtc tgggttgaggg tctggtcctg gaccacggtt 660
ctcgtcacc tgacatgaag cgcagggcag agaattgtta catcttgacc tgcaatgtct 720
cattagaata tgagaaaagt gaaatcaatg caggattttt ttactcaaac gcagaacaaa 780
aaaaaaaaa agg 793

```

```

<210> 139
<211> 526
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:897616.1:2000SEP08

```

```

<400> 139
gcgggattgg ctccgggggc gcggcgagga gttttttggt tgggggtggg gtggcctgtg 60
cctacagaga catgaggctg agctggatcc tgacagtact gtccatttgc ctgagtgcgc 120
tggtcacagc cacaggggcc gaaggcaagc ggaagctgca gattggagtg aagaaacgtg 180
tggatcactg tcccatcaag tctagaaagg gagatgtctt acacatgcat tacacgggaa 240
agctagaaga tgggacggag tttgacagca gcctaccaca gaaccagccc tttgttttct 300
cccttggcac tggccagggtc atcaagggct gggaccaggg gctactgggg atgtgtgaag 360
gggaaaagcg gaagctagtg atcccgtctg agctggggta tggagagcgg ggagcccccc 420
caaagatacc aggtggagca accctgggtg ttgagggtga actgctcaag attgagagac 480
gttcagaact gtagcaggct gcagaggggt aggaggcccc tatcag 526

```

```

<210> 140
<211> 431
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:736860.1:2000SEP08

```

```

<400> 140
gttcgccatc ggcagccgag caccatgagc tccacaggcg gcggcgacga tggcgccaag 60
aagtcttggc cggaagtggg cgggctcagc ctggaagaag ccaagagggt gatcctgtgc 120

```

gacaagcccg acgccgacat cgtcgtgctg cccgtcggca cgccgggtgac catggatttc 180  
cgccccaacc ggcgtccgcat cttcgtcgac accgtcgcgg aggcacccca catcggttga 240  
ggtttaaatac tacaaaatga atgactcgga catgccatgc gtacctgtcc gtcgccgaat 300  
aatggatgtg tgtgtgcttc gatcgttcct ataagttgct agtaaaaata atggcatcgt 360  
cgttatgcat gaataaaaag tatcagaata atgttcaccc tttcacgatt ggataaaaca 420  
aaaaaaaaag g 431

<210> 141

<211> 1429

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:027066.6:2000SEP08

<220>

<221> unsure

<222> 225, 260, 275, 301, 307, 1351, 1424

<223> a, t, c, g, or other

<400> 141

cttttgccctt ggggtgtttct ctccctgccgt cccgtccaca cgcttcccg gttccctgcc 60  
gcccgaggga ttgccacagg gaagtaccac gcccggtgac tcaccaacag cgctgagtgg 120  
gaggccgcct gtgtgaaggc gggcaggaag tgtggggacc tgggtgcacc gctggtctac 180  
tgccccgagc tgcacttcag cgagttcacc tcagctgtgg cgganatgaa gaactcagtg 240  
gcggtaggtt tggagcctcn aacctggagc ctgcnacatg ggtggaccgg ggcaggcgga 300  
ncctgcnttc aggggtgtggt gcacccaggg agctggggcc cccagaagc agccacagtg 360  
cagacgaggg cttgagaggc aggcgtcagg gcacaggagt catccagaca gcgtggggcca 420  
ctcactggct tccctgccac acagccaagg gtttctcccc cagtcttggg tctggctcag 480  
ttgccccatc aggcctcttg ctggcttccc ccttggccta tgggtggggc agactcctta 540  
gctcatggtc aaggccctcc cagcccagct tctgctgctt cccacacgc tccctcccca 600  
cacgctccct cccagccacc ccgagctcct tgacagacagc aatagtgaca ggcgatgggg 660  
cagggtggag agggccccgc cggagcaacc cacaggcact gtgtcctcct ggcctcccta 720  
ggaccgagac aacagcccca gtcctgtgc tggcctcttc atcgccctcac acatcggctt 780  
cgactggccc ggagtctggg tccacctgga cattgctgca ccggtgcatg ctgtgagtgt 840  
ctccccctcc cactggccct ggctgctccc gcccgcttgt ccaaacagcg cccctctggc 900  
tctggagctg ctggcagagc tcatcagaaa cttctgtctg tgacccagct tccagcccgc 960  
tgtccccacc acccccagggt ctcatcctcc ctgggaacag agtggctgct gtgtgagacc 1020  
cttccccagc cagcctgtcc tccatagggg atcctggggc ctgtctcacc catccccacc 1080  
ctgaggagct cccgggggtga aggcagagca cacaggccct tgccccctgc ctacgcctgg 1140  
cctgccagcc ctgaacgtgt ccagccagca gcatggaagg ctctgggctc cggctggtgc 1200  
tcaggatctc ctctctgaga aggggactgt ggggcacgtg gaggggaccc aggaggtgag 1260  
gggtccccag gaacccctcc tgtgctgcag cccacgcgcc agagtctgtg tccctgccctt 1320  
tgcttgacag gtgagcgagc cacaggcttc ngtgtggccc tccctgtggc gctcttcggc 1380  
cgtgcctctg aggacctctt gctgaacctg gtgtccccac tggncctgtg 1429

<210> 142

<211> 663

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1074263.1:2000SEP08

<400> 142

gcggaacccat ggccagcccg ctgcgctcct tgatgctact gctggccgct ctggccgtgg 60  
cctggggccgg aacctccagg ccacccccgc gattgttggg agctccgcag gaggcagatg 120  
ccagcgagga gggcggtgcag cgagcggttg acttcgccgt aagcgagtac aacaagggca 180  
gcaacgatgc gtaccacagc cgcgccatag aggtgggtgag agctcgtaag cagcttgggtg 240  
ctggataaaa ctattatttg gatgtggaga tggggccgaac tacatgtacc aagtcaccga 300  
caaatttgac taactgtcct ttccacgacc agccccatct gatgaggaag gcactctgct 360

```
ccttccagat ctacagcgtg ccctggaaag gcacacacac cctgacaaaa tccagctgca 420
aaaatgccta agagctgagt ctcataggac catgccaatg gtcccttact tgttccccta 480
ccctgtagtg ttttatccct gagaagggtg ctccagctct ggagggcatc tccgggggtg 540
tcccaccagg agacagtaaa gaagctgctg caggcaggtt ctgcacgtca gaacagctgt 600
cccttggttc tcttctcctt gcagtacctg tcataccttg ctcttgctca attaaaaaat 660
ttc 663
```

&lt;210&gt; 143

&lt;211&gt; 701

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:334345.1:2000SEP08

&lt;400&gt; 143

```
gagcgcggga agtttgccgg aaacggacag aaaggagtcc ctgcgaccgc gtagctcctg 60
cagcgccccc tgttggcaga gtcgtagcga ggccgagatg aggactctga cccagtgca 120
gtgtctgtag ggaagaggag ccatggggct tcgggcaggc cccatcctgc ttctgctgct 180
gtggctgctg ccaggggccc attgggatgt gctgccttca gaatgcggcc actccaagga 240
ggcggggagg attgtgggag gccaagacac ccaggaagga cgctggacgt ggcagggttg 300
cctgtcgttg acctcagtgg ggcattgtat tgggggctcc ctcatccacc cacgctgggt 360
gctcacagcc cccactgct tcctgagggt gactccgggg ggccgctggt ctgccccatc 420
aatgatacgt ggatccaggc cggcattgtg agctggggat tcggctgtgc ccggcctttc 480
cggcctggtg tctacaccca ggtgctaagc tacacagact ggattcagag aaccctggct 540
gaatctcact caggcatgtc tggggcccg cagggtgccc caggatccca ctccaggcacc 600
tccagatccc acccagtgct gctgcttgag ctgttgaccg tatgcttgct tgggtccctg 660
tgaacatga gccatggagt ccgggatccc ctttctggta g 701
```

&lt;210&gt; 144

&lt;211&gt; 920

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1093914.1:2000SEP08

&lt;400&gt; 144

```
gaaaaattac acaatctcag cctgggtaaa tgagagatgc ccttttagct tattctcccc 60
agttcaccct ttctcccag gtgattgagt gtggatcaag gactggaacg tagcctcttt 120
gtgcccacag tggaaaggac ccagattgt cgtcctgacc actcccaccg ctgtgaaggt 180
agaggggaatc ccagcctgga tccaccacag ccatgtaaaa cctgcagcac ctgaaaccta 240
gaaggcaaga ccaagcctgg acaacccttg cagagtgacc ctgaagaaga caagccctgc 300
tccagtcaca cccggaagct gactgggtcca cgcattggccg aagcatgagg aagttcactg 360
tggaactcat ttttttaaat tttggacttg tacaagtaaa ggacttcaac tgaccttctc 420
cagactgata actgtttcca gtatatacat caagtcaactg aggtaggaca aaagattgct 480
acagtcttat tattttatgg ttattataag tgtaccggga ctctaaaaga aacttgtctg 540
tataatgcta ttctatccaa ggtatgcagc ccagggaata accaacctga tgcgtattat 600
gacccatttt aagcctccca tggtcacagt ttttaaaata aaattaagga ctggctcctt 660
tctaggtgac acaagtaagg taatagctag aacagaagaa agaaagggtcc ccaaaaatgt 720
aaccttaaaa tttgacgcct gtgccgctat taatagtaaa cagcatggga taggatgcgg 780
ttctctagat tgaaaaaaa gttacacagc agaaaataag tacatctgtc aaaaatcata 840
tttatgtgag atgtgtcaat actagtcttg tgtcatttag gctacttaga aagaagataa 900
aaaatatact gtttggtccc 920
```

&lt;210&gt; 145

&lt;211&gt; 2762

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;



&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1188168.1:2000SEP08

&lt;400&gt; 145

```
ggcgtactgg cctgggtccag cacctgcggg gccctcgtgg cttggaggggc tggggccgggc 60
ggggaacggg cggggcgggc cggaggcggg ggcggctgac tcgccttctc tccggggctg 120
cgaccccgag gcaaccggct cgagatgggg agcccgcgga gccaggatg cggggcgggc 180
ggggcgcgac gccggctgag ggagctgttc cgggacgccc ccttccccgc cgcggactcc 240
tcgctcttct gcgacttgct tacgccgctg gccagttcc gcgaggacat cacgtgtaag 300
gctgtgcccc ggaagatttt gtgccacacc ccggctgttt ccaggatgac ccactgggaa 360
gggcagtgtg aagcagggtg ctgctggggg attgctgggt cctgtgtgac tgcgccgcgc 420
tgacagaagag acaggcacct ccctggacca ggtcatctcc ccgggacagc cgagctgggc 480
cgaccaggag taccggggct ccttcacctg tcgcattttg gcagtttgga cgctgggtgg 540
aggtgaccac aagtgaccgc ctgccgtgcc ttgcaggag actctgtttc tcccgtgcc 600
agaggggagg tgtgttctgg ctccccttac tggaaaagg ctacgccaa gtccatgggt 660
cctacgagca cctgtgtggc ccgggcagtg tggcgatgc cctggtggac ctgaccggcg 720
gcctggcacg aaagactgtg aaccctgaag ggcgtagca ggaagcggag gccagtcagg 780
acagtgccta ggtgccgctg ggagcacagg acttgtcgcc aggtgtctcc acctgaagga 840
tccagtgtct gatcagctgc ttgcgtgctc agccccagta gcaggtgcgc ccggggagct 900
gggggagttc catgccttca ttctgtcgg acctgtcggg agctccaggg ctccaggcgg 960
ccagtgcata ctgctgcttg cggatccaga acccctgggt tccgtcgggt gctggcaggg 1020
gctctggaga gaggggggtg aagggtggag ccaggtagat gcagcggtag catctgagct 1080
ccctgttcca gctccaggaa ggggatgttc tgggtggagg aggaggagtt cctcaggagg 1140
tttgacgagc tcaccgttgg ctaccgggtc acggaggccg gccacctgca gaagcctcta 1200
caccatgatg acgtgtctg ccatacgcgg gcgctgcctg gggcctgggt caaggggcag 1260
tcagcaggag gctgcdggaa caacagcgcc ttttcccagc aaccccaaat tcttgcttcg 1320
ggtctcagaa ccgagtggag tgtacatttc cgtcctgcag agatccaggc tgcacgcggc 1380
ggactgggca ggcggggccc gggcactggt ggttgacagt catacttcgt ggagcccagc 1440
gagcatcccc ggcaaagcac taccaggctg tgggtctgca ccctctggaa ggtagagaag 1500
cggcgggtca atctgcctag ggtcctgtcc atgcccccg tggctggcac cgcgtgccat 1560
tcaatacgac cgggagggtc acctgcgttg tgagctctca ccgggctact acctaggctg 1620
tccccaggca cctatcctag aagcgaacgc gccacgggga cgttctgct ccgagtcttc 1680
tctaccgggc gagtctccct tagcgccatc agggcagtgg ccaagaacac cgcccccg 1740
ggcaaccctt gccttcgggg ggagtgggg accgtgcagc tacggggttc ttgaagagtc 1800
ggccagacgg cgggtggggca gcaggaaact tgccttcata acccagcaa ccctggttcc 1860
cccttcttgg gttccccgag gggccctgtg gccccgtct agcgtccgca tcaactctga 1920
tcaagcactg cctggtccca agtagacacc tgaagttcca cccaatcgg gcttccaata 1980
tcttccgagt gtccccagag ggtgggacag gagccatgga cgcacccac actgctggct 2040
ggcaggaagc cagctagctg agctggcggt gccacatccc gctacgcccc aggagggtga 2100
gcccggctcc tgcttctct cctgcggcgc acctaacaag gttgtgccct ccacctacc 2160
tgccccgaca cagagggggc cttcacagtg accatctgca accaggattg acgaggccat 2220
cccattcaca gccaaaggaga tgctggggcc agttcctcca aagagggtct catcataggc 2280
agtgatgatg aacctaacag ggtgggcccc ctgggtgccag ctccagggtga ttggagcccc 2340
gagggcctga cacggttctc cagcacgtct gggaccggcc agcccttgca cctgtggggg 2400
ctggtccatg agtcttggcc ttgccttccc agcccttgcc aggggagact gcggacttag 2460
gggtccacgg gaagcgctcg tcaaggagag aacgcagccc tggggggcca gtctggtgca 2520
tgcaaggaag ggtgggaagc ttgcctggct tctgttgcgc cacttagacg tgccgatgac 2580
cccaggatcc cacaacttcc caggatccct tccaaaatcc tttgtgact ccataatgga 2640
ggcctcaaca cccagagggt agggcagcag atcttcttta taactattta tggttcgaga 2700
tcacgtttta ggggatgtaa ctttataaat aaacatgagc gcggatgatt tgcaaaaaaa 2760
aa
```

&lt;210&gt; 146

&lt;211&gt; 584

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1065168.1:2000SEP08

&lt;400&gt; 146

```
gcgatgcttt ccactgccgc ttaccgggac ccggaccgag agctggtaat ggggcctcag 60
```

```

gggctccgcc ggccctgtgc agatgcgttt ctcgccttat gccttcaacg gaggtactgt 120
attggcaatt gctggagaag atttttccat cgttgcttca gacactcggg tgagtgaagg 180
gttttcaatt cataccgag atagcccaa atgttacaaa ctaacagaca taaactagta 240
attggctgca gtggctttca tggagattgt ctcaccttga caaagattat tgaagcaaga 300
ttaagatgt acaagcattc caataacaag gccatgacaa caggggcatg tgctgcaatg 360
ctgtctacca tcctgtactc acggcggttc ttcccttact atgtctacaa cattcaattt 420
aggaaggaca ttgatgaaga aggaaaaggg aagaccgtgt acagctttga acccaagtag 480
gactacttat ccagagagaa ctctattcaa aggacacgga ggactcagca atgtgaccaa 540
tagacatgga caagacacga ctgactcgac aaccaaggat taga 584

```

&lt;210&gt; 147

&lt;211&gt; 1065

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1180418.1:2000SEP08

&lt;400&gt; 147

```

gaaagaaaga aatgctaggg gaaaatgttt taactagtca ttcttcccag tagctaataga 60
agctgacttt taaaaagaag gctgtgagct ttgcagatgc tgctgccgcc cagggccccc 120
tgcttccagc catggtcaac cccaccatgt ttttccacat tgctgtcgat ggcgagccct 180
tgggctgtgt ctcttccgag gtaaggggcc tggaaagcaa gaagtgactg ctcatctaata 240
ccataaaact atgttaacag attggagctg tttgcagaca aggttccaaa gacagcagaa 300
aatttccatg ctctgagcac tggagaaaaa ggatttggtt ataaggggtc ctgctttcac 360
agaattattc caggggtttac gtgtcagagt ggtgacttca cagccatac agcattgggtg 420
gcaagtccat ctgcagggag aaatttgatg acaagaactt catcctgaag catacgggtc 480
ctggcatctt gtccatggca aatgctggac ccagcgtgaa cgtttcccag ttttttatct 540
gccctgccaa gatgccaaga cagagtgggt ggattgcaag catgtggtct ttggcaagg 600
gaaagatggc atgaatattg tggaggtcat ggagcacttg gggccaaga atggcaagat 660
cagcaatcag caagaagatc accattgctg actggacaac tgcaataaat ttgacgggtg 720
tttctcttaa aacaaaaaaa aaaatactgt gacagaccaa ggtaaattgt ttttgataat 780
tacagaagtg tgatttctta atcagcagct gtcaaacata gtgttcctta atttaaaagc 840
ttgaacacta aattataaat tggagaggtt ggaataatta cgtcatatc tctagaaaca 900
caagtcttta gtagcaaaaa agaaattagc aagagaagaa aactgttcag tactttgaaa 960
ggaaaaagtt ttcagtata gtttttttag atgaaaatta acatgataaa gaagtgcag 1020
gccaggtgtg gtggttgacg cttttaacac caacaatttg ggagg 1065

```

&lt;210&gt; 148

&lt;211&gt; 2152

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:232648.1:2000SEP08

&lt;400&gt; 148

```

aggcagggcc cttcagggtca gcccgagca ggctgactcc ttgactcccc agctgacgcc 60
ctgggggactg atctcgggga ctgaggatgc aggtgcacag tagggctgcg gaccgggctc 120
gggtggcttcg caccggagcc cgggtggggc ggtgcccgtta cctctgtatg ctgggctggg 180
ggcgaccac gtgctcacag cacacgcacc cagcgtggc tcggcaggcc ggccatggag 240
ccaggcagcg tggagaacct gtccatcgtg tacggagcc gcgacttcct ggtggtcaac 300
aagcactggg acgttcgcac tgacagcaag gcgtggcggg agactctgac cctgcagaag 360
cagctgcggg acgcgtttcc cgagctggcc gaccctgaca cctgctacgg gttcagggtc 420
tgccaccagc tggatttctc caccagcggg gcgctgtgcg tggccctaaa caaggcagcc 480
gccggcagcg cgtacagggtg cttcaaggag cggcgcgta ccaaggctta cctggcattg 540
ctgccccggc acatccagga gagccgggta accatcagcc atgccattg caggaacagc 600
acggaggggc gggccacac catgtgcac gagggctcgc aggggttgtga gaacccaaag 660
ccaagcctca cagatctcgt ggttctggaa cagggctgt acgcaggcga tctgtctcc 720
aaagtgtcgc tgaagccgct cacgggccgg acacaccagc tgcgcgtgca ctgcagtgc 780
ctggggccacc ccgtgggtggg cgacctgacc tacggagaag tctcgggccg ggaggaccgg 840

```

```

ccgttcagaa tgatgctgca cgctttctac ctgcgcatec ccacggacac cgagtgtgtg 900
gaggtctgca cgctgaccc cttctgccc tccttgatg cctgctggag cccccacaca 960
ctgctgcagt cgctggacca gctcgtgcag gccttacggg ccacaccga ccctgacccc 1020
gaggataggg gccccaggcc aggcagcccc tccgcactcc tgcttggggc cggccggcct 1080
cctccacccc caaccaagcc cctgagact gaggcacagc gggggccctg cctgcagtgg 1140
ctgtcggagt ggacgctgga accggacagc tgagagccgt ggggctgggg cagggggtgt 1200
cagctgcaca gcgggactct agggagatgg gcgagcgagc gtctgtcac tggctctggg 1260
gcctcgaggt gccaggcagc atcaggccca ctgggttgcc ccggccaggc ctgcgaggaa 1320
gggctgaggt ggggcccggca gggggcgcca ggagccgtg atcacagggt acgaccgcac 1380
cgcgcccggt ggactgatgc gggatcccga ggggccttcc tgcccatatg ccccgggaga 1440
aaccgaggcc cctccctcct cctggaacag cttccggctc tcaagcgtca ccccgagggc 1500
gtcagtttta cggactcaag gtcacctcag gaagaggcag gggcagggtt tgggataggc 1560
tttgccctca ggaagggtg ctctggggc ttgtgagcta ctgccccaa cctacctct 1620
agaggggctg ggaaggggcg tctggggctc acctggcctg ggagaccat ctggctcctg 1680
cgtcctctgc cctcactgc tctgtgcaga tctgtgcgc ctcagctgcc tcctcccag 1740
acctaatggt cctgctggg ctgagctctg caggcccgcc tgctgtgccc ttggcctcac 1800
tgtaccagt gttccctctc tgcccggatt ctgagctcag tgtggtgtt ggtgcacagg 1860
ggttggtcag gggccatggc caaggccctg ccacgcacgc ccatccctca gatccactgt 1920
gagcaccaac ctgctgcagt ctcttgggcc cctgctggca gctctgccac gtcaccgct 1980
gcctggctcc cacacagcca tgcatgtca ctctgcctcc gggaccccag cttgggagct 2040
gtgggtctgc caggtccac ctctctgtc cccatgcca caacctgggc tcctggctac 2100
agcagggtcc cagggactcc aaataaatgt tcagtgaact gctccaagga aa 2152

```

<210> 149  
 <211> 972  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1078420.1:2000SEP08

```

<400> 149
cccaccttac ctcatgccc gctggggcgt gttcacccct ccgcgctgtc acaatgcct 60
tgccctccct tgctctctg tctactctgc gttccccgt gggaggaggg tctggtgata 120
aaataggggc gagactatgg aaatgatgtt agacaaaaag caaatcgag caattttctt 180
attcaagtcc aaaatgggtc ataaagcagc ggagacaact cacaacatca gcaacacatt 240
tgggccagga actgctaatt aacgcacagt gcagtgggtg ttcaagaagt ttcacaaagg 300
agaggagagc cttgaagatg aggagcatag tggccagcca tcggaagttg acagtgacca 360
attgagagca atcatcgaag ctgacccctc tataactaca cctgaagttg ctgaagaact 420
caacattgac cattctacag tcgttcggca tttgaagcaa attggaaagg tgaaaaagct 480
tggttaagtg gtgccttggt aactaagcga aaaacaaaaa aattgtcgtt ttgaagtgtc 540
atcttctctt attctacgca acaacagtga accatttctc gatcagattg tgatgtgcca 600
tgaaaagtgg attttatatg acagccagtg acagccatag tcagtgggtg gtaccaggaa 660
gctccaaagc atttcccata agccaaactt gcaccagaat aaggtoatgg tcaactgtttg 720
gtggtctgct gctggtctga tccactacat actgttctga atcctggcgt taacccttac 780
tatctgaga aatatgcttc agtaaatcga atgtagatgc taccataaat actactatac 840
acctgcccgt ggctattggt caactagaat agggcccaat tcttctccac gactagtggc 900
ctaacctact acattgcaca accagcactt gggaaaattg aacgaattgg gctacgaagt 960
tttgctcat cc 972

```

<210> 150  
 <211> 696  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1397599.1:2000SEP08

```

<400> 150
ctgttttctt ttgaggaata ttgatgatcc cactgcttga gtaatggaaa ccagcttgca 60
gctgatagtt attttttgtc ttttttctg tctatocac tgccttaggt ttggaatttc 120

```

```

agctgggcag cttcgatccc accacctgta gccactgagg ttcacaggct ctgtttgagg 180
tggtgtgtta gattgtgctg ggttggtttt gctgcttatt gaggatgaa gagcaacata 240
ttagctactg agataatttg cagaaaaagg aagagatgtc ttcagctcaa atgctatcaa 300
acttgtgaat gtcacacgcg tccatgtgaa gagactacca aacaggcttt gtgtgagcaa 360
caacgctggt tattttggtg ctgaagaccc gggacaggag gactccttca ggagacgggt 420
cccctgtcct tgccctcact ccgtgaggag atccacctac gacctcgggt cctcagacca 480
accagcccaa ggaacatctc atgaattttca aatcggaattc ccaactatat gaagacaccc 540
tagctggacg atcagttctt attaagaacc tgacctctca aactctacaa cctcgatgga 600
ctggacccta cttagtcatc tatagtaccc caactgccgt ccgcctgcag gatcctcccc 660
actgggttca ccgttccaga ataaagctgt gcccat 696

```

&lt;210&gt; 151

&lt;211&gt; 977

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1397655.2:2000SEP08

&lt;400&gt; 151

```

gggtggcatt ccttggccct taggtacgcc agtggtcacg tgtttaagtt gtctgcgtcc 60
ccatcccagtg tagccccag gagcaggggc ttggtttggt cactgctggt caccactgc 120
tgaaaataac accccctttg tcaaagtgga gtaagctcga gccgtattgt taattcatag 180
cttctgtttt tgtcgcccag gctagaattc agtggcagga tcatagctca ctgcagcctc 240
gaactcctag gctcaagcaa tctcctgct tctgcctcct gggtagctgg aactgtagat 300
gcaagccacc aagcccagct ctaatcacac tttgaataaa aagtatggat ggaggaccag 360
cagcatctgt atcacctggt tggaaatttca gctgggcagc ttcgatccca ccacctgtag 420
ccactgaggt tcacagggctc tgtttgaggt gttgtgttag attgtgctgg gttggttttg 480
ctgcttattg atgatgaag agcaacatat tagctactga gatattttgc agaaaaagga 540
agagatgtct tcagctcaaa tgctatcaaa ctggtgaatg tcacacgcgt ccatgtgaag 600
agactaccaa acaggctttg tgtgagcaac aacgctggtt attttggtgc tgaagaccgc 660
ggacaggagg actccttcag gagacgggtc cctctgcctt gccctcactc cgtgaggaga 720
tccacctacg acctcagggtc ctgagaccaa ccagcccaag gaacatctca tgaatttcaa 780
atcggattcc caactatatg aagacaccct agctggacga tcagttctta ttaagaacct 840
gaccctcaa actctacaac ctcgatggac tggacctac ttagtcatct atagtacccc 900
aactgccgtc cgcctgcagg atcctcccca ctgggttcac cgttccagaa taaagctgtg 960
cccatcgga aaaaaaa 977

```

&lt;210&gt; 152

&lt;211&gt; 1265

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:241055.1:2000SEP08

&lt;400&gt; 152

```

gttattatta ggttcaatcc tatcatctac tatttatttt atatttgatc tttgttccca 60
ttctgctgta tttttctgcc ttttttgagt taatattagg ttagtgcaaa tgtaattata 120
gcttttgcat tgttagaatt tgctgtttga cattggaata cgttcttaaa taaatgtggt 180
taggttatac atcattttaa tgggcattgc ttgctttatg gttttttgct aatgacttct 240
tacttgctgt ttatttttaca tttatttttg actatggaaa tgatgttaga caaaaagcaa 300
attcgagcaa ttttcttatt caagttcaaa atgggtcata aagcagcgga gacaactcac 360
aacatcagca acacatttgg ccaggaact gctaataaac gcacagtga gtggtgggttc 420
aagaagtttt gcaaaggaga ggagagcctt gaagatgagg agcatagtgg ccagccatcg 480
gaagttgaca gtgaccaatt gagagcaatc atcgaagctg atcctcttat aactacacct 540
gaagttgctg aagaactcaa cgtaagaact caatgttgac cattctacag ttgtttggga 600
tttgaagcaa attggaagg tgaaaaaact caattaagt gatgcctcat gagctgagt 660
aaaattttaa aaatcgcaag ttttgaagt tcatctctc ttattctaca caacaataaa 720
gaaccatttc ttgatcggt tgtgacatgt gacgaaaagt ggattttata cagcaactgg 780
tgacgaccag ctgagtggt ggaccaagaa gaagttccaa agcacttccc aaagccaaac 840

```

```
ttgcacacaaa aaaagggtcat ggtcactgtt tgggtggtctg ctgccggtct gaaccactac 900
agcttttctga atcccgggtga aaccattaca tctgaaaagt atgttcagca aatagatgcg 960
atgcaatgcc tacagccagc actgggtcaac agaaagggcc caattcttct tcacaatgtc 1020
ttgactgcac attgcacaac caatgtttca aaagttgaac aaactgggct acaaagtttt 1080
gcctcatcca ccatattcac ctgactttct gccaacccgac taccacttct tcaagcattt 1140
tgacaacttt ttgcagggaa agcacttcca caaccagcgg gatgcagaaa atgctttcca 1200
agacttcttc aaatcccga gcatggattt ttacgttata gaaataaact tatttctcat 1260
tggtg 1265
```

&lt;210&gt; 153

&lt;211&gt; 924

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1101065.1:2000SEP08

&lt;400&gt; 153

```
aaatatggtg gcaaacctgt gtgtattgtg tgtattctga ttgttccacc aactggctgt 60
tcttccatcc ctctccctct ccttgagcca cctgttctc tgaggcacia caatattgaa 120
attggggcaa ttaataaccc tacaatgact tctaagtgtt caaatgaaag gaagagcagc 180
acataaaaagt ctcacattaa atcaacagct acgaatgaca actttgtgag aaaggcatgt 240
tgaaagctga gacaggtcaa aagctaggcc tcaggtgccc attaagcatg ctgtgaatgc 300
aaaggaaaag cttaaaaaaa tttttttaag ggtgactcca gtaaacacat gaatgataaa 360
gaaagcaaga caggattatt gctgatattg agaaagtctc agtgatctgg gaacatcact 420
gaagatcaaa ccagccacaa gcagtcctct aaggccaaag cctaattccag agcaaaaggcc 480
ctaactcttt tcaattctat gaagactgag agtgggtgag gcagctgcag aagaaaaagg 540
tgaaagaggt gaggaagctg cagaagaaaa cctggaagtt agcagaggtt ggttcattag 600
gttttaagga aagaagtcac ctccatatca taaaaagtgc aacgcgaaag cagcaagtgc 660
tgatggggaa gctgcagcaa gtcattcaga agatcttgct aagggtatgc acagatgtgg 720
aaacaggaac tgatgtgtcc attacaccac taggacagag gccagaacaa tgaagaaacc 780
aaatacttgg aagagggtag agataatgaa tggagtccaa gagccctgat tgtgccataa 840
atgtccagat aattccatac ctgaggatta tgtggtttgt aaacttggca cttagaagaa 900
ccaataaaat catgttatag tttc 924
```

&lt;210&gt; 154

&lt;211&gt; 585

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:475629.1:2000SEP08

&lt;400&gt; 154

```
ctcgcagacg cgcctaggct tagctagggt ctgggtctgc tgccccgccg cgcgcgcatt 60
cacaagcttc tagacgttcc gaggaggagg ggaggtaccc gccgccgcag ccatggcgctc 120
gactaaggctc cagcgtatca tgaccagacc catcaacctc atcttccggt tcttacagag 180
caaagcgcgc atccagatct ggctcttcga gcagaaggac ctccggattg agggctcgcat 240
tatcggaattc gatgagtaca tgaacttggt cctcgaagac gctgaggaaa tcaacgtcaa 300
gaaaaacact aggaaatcac tgggccggat actcctcaa ggtgacaata taactttgat 360
gatgaacagt ggcaagtga gattcttatg gcaggaacga tgtttggtgc tcgcttcgat 420
tatctgtgac acaatgtact ccatccatcc gtccagactt aaatgatgta tcatgtcatt 480
gtgggtttag gagagttgta taactacctg aggaaactca cgtgttacta tatgctaaga 540
tgctgtgacg gtgagctatt cagtgtcac ttgtgttcca aagat 585
```

&lt;210&gt; 155

&lt;211&gt; 938

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:348991.1:2000SEP08

&lt;400&gt; 155

```
ggccggggcc gggggcgcg gggcgggcgaa ggcaaggatc gggcgggcgg tggacagcgg 60
cgggcgggcg tggccaagag caagagccgc cgaggaagg gcgccatggt ggtgtcgggt 120
gagccgcacc ggcacgaggg cgtcttcac taccgcgggg cgaggagcgc gctggtcacg 180
ctgaacatgg tgccggggcca gtctgtgtac ggcgagaggc gcgtcacggt gaccgagggc 240
ggcgtgaagc aggagtaccg cagctggaac ccgttcgct ctaagctggc cgcggccatc 300
ctggcggggg tggaccagat ccacatcaag cccaagtcca aggtgctgta cctggggcgcc 360
gcgtcgggca ccaccgtctc ccatgtctcc gacatcattg gccagacgg cctggtctac 420
gccgtcgatt tctcccaccg cgcgggccc gâtctggtca acgtggccaa gaagcgcacc 480
aacatcattc cggtcctgga ggacgcgcgg caccgcgtca agtaccgcat gctcatcggg 540
atgggtggac tgatcttcgc cgacgtggcc cagccggacc agtcccgcac cgtggccctg 600
aacgcccaca ccttcctgcg caatgggggc cactttctca tctccatcaa ggccaactgc 660
atcgactcca ccgcacccgc cgaggtctgt tttgcttctg aggtgaggaa gttgcagcag 720
gagaaactga agcctcaaga gcagctgacc ctggagccct atgagcggga ccacgctgtg 780
gtggtcgggg tctaccgacc tcttcccaag agcagcagca aatagcacc agctcaggct 840
cgccctgcat ctccccaagg ctgcattgtg tttgttatta ttttctgtgt gttttctttg 900
tgagtgtttt gttttgttgt ttttctatta aactgcac 938
```

&lt;210&gt; 156

&lt;211&gt; 588

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:475629.1:2000SEP08

&lt;400&gt; 156

```
ctcgcagacg cgcctaggct tagctagggt ctggtctcgc tgccccgcgc cgcgcgcatt 60
cacaagcttc tagacgttcc gaggaggatg gggagggtacc cgcgcgcgc gccatggcgt 120
cgactaaggc ccagcgtatc atgacccagc ccatcaacct catcttcggt ttcctacaga 180
gcaaagcgcg catccagatc tggctcttcg agcagaagga cctccggatt gagggctcga 240
ttatcggaat cgatgagtag atgaacttgg tcctcgaaga cgctgaggaa atcaacagtc 300
aagaaataac actaggaaat cactggggcg gatactcctc aaagggtgaca atataacttt 360
gatgatgaac agtggcaagt gaagattctt atggcaggaa cgatgtttgg tgctcgttc 420
gattatctgt gacacaatgt actccatcca tccgtccaga cttaaagat gtatcatgtc 480
attgtggttt agtgagagtt gtataactac ctgaggaaac tcacgtgtta ctatatgcta 540
agatgctgtg acggtgagct attcagtgct cacttgtgtt ccaaagat 588
```

&lt;210&gt; 157

&lt;211&gt; 451

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:261331.1:2000SEP08

&lt;400&gt; 157

```
gaccattttc cagtaacatc ttcctggcaa accaccaagg gacaacactg gagaaaccpc 60
gacccaaaag aatagactgc agcactaatt ggccaacttt gcactacgga aatgacatta 120
gacaaaaagc aaatttgagc gattttcgtg tcaaagttcg aaatgggtca taaagcagtg 180
gagacaactc gcaacatcca caatgcattt ggcgaggaa ctgctaata actgacagtg 240
caatgggtgt tcaagaagtt ttgcaaagga gacaagagcc ttgaagatga ggagcacggt 300
ggccagccat cagaagttga taacgaccag ttgaaagcga tcgtcaaggc tgatcctctt 360
acaactacac gagaagttgc caaagagctc aacatcaacc attctgtggt cattcagcat 420
ttgaagcaaa tggaaggtaa aaaagcatgg t 451
```

&lt;210&gt; 158

&lt;211&gt; 1839

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:815686.1:2000SEP08

&lt;400&gt; 158

```
gaaaaagctgg tattccttta taagatctag aagttaaggg tctctttata tgtgtagtaa 60
tccctcaagt tgcctaagat cattttaagg tcaagctggc ctaatatatt caatgagata 120
aaacttagtt ctttatctcc tctcaaacag aagaaaaatg ttttgTTTT accatgggta 180
caaatagata ctggTTTTTT ttttcaaata attcaagttt ctaaacccta gcctagccta 240
tctttctttg ccattttatac tggttgcctt gaaatgaggg gaactctctt acccctgaga 300
ataaccagtt aaccctcca gttctggctc agtggtatat gaggggactt cagaaagttt 360
gtggaaaact ggaattaaaa gataaaaaaa tatatatata aactttaatt tattgacata 420
agtgccacca gctatcaagt tcaagacact ttggttaagca aaaaggtgat ataccagcat 480
ttaatgcatc cctaaagaac tgagggtcct gggaatttaa ccatgtcaat gtagtctttt 540
ttatattatt aactaaagaa caatgggtac cttttacaaa gtttttaaga tcaggaaaca 600
aaaaaaagga ggagccaaat aaggactgta acatgagtgc ctatggattt ccattggaag 660
ctctcacaaa attgcccttg tttgatgaga ggaataagca gaagccttg tgatggggaa 720
ggactctcta atgaagcttt ccaggtgtt tttctgctaa agctttggct ttctcaaaac 780
actttcataa taagcagatg ttgttgttct ttggccctcc agaaaaccaa ctagaaaaat 840
tcctttaaca ttccaaaaga ctattgccat gagcttggtt cttgaccagt ccagttttgc 900
tttgactaga ccgcttcac ttcttggtag ccaattgctt tgaattggtt ctttatcttt 960
aggatcaata ctggtaaagc catgttttgt ctctgttaa aattcttcaa gaaattgcta 1020
caggatcttg atcctacttg ttaaaatttc cattgaaagc tcagctcttg tctgcagctg 1080
atctggggcac aacagttttg gtacccattg agtggaaggt ttgctcaact ttaatttttc 1140
agtcagaata gtgtaagcta agccaattta gatatctatg gtgttgacta ttgtttgtgc 1200
tgtaattctt tgtcttctt tcaattaagg catggacata atttttttca atctttttgt 1260
ttcaaacgat tggaaggct tgctgctatg tgcttcatct tcagcatcat ctcatccctt 1320
cttaaaatga gttatctgtt tgcaaaactgc tggatttctt ttggagaatc atcccccata 1380
aacttctcat aaagcatcaa tgattttacc attcttccac ccaagcttca cactaaattt 1440
gatgttggtt gctcaaaaat ttttatagta gaatttatgt tgctctgaca ggtgctcttt 1500
tcaaacttac gtcttacctt tttagtgcct caaactacat cctgttcagg catgttataa 1560
caagttagta tgagttttatt ttggagcaaa agatttgaaa tccatgaata gttttattca 1620
taagattcat tttccataaa cttttgaaaa cacctcataa aaatagctga ctcttgaaaca 1680
acacaggttt aaactgggtg gggtgccact tatatgtgga tttttttcag taaatttgtt 1740
ggaaaaatgt ttgaagattt gcaacaattt gaaaaactcg cagctgaaac acataaccag 1800
aaaacacatt gacttcagac aatctgccag gagagtcc 1839
```

&lt;210&gt; 159

&lt;211&gt; 678

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1167327.2:2000SEP08

&lt;400&gt; 159

```
aactgttttc ctttgaggaa tattgatgat cccactgctt gagtaatgga aaccagcttg 60
cagctgatag ttattttttg tcttttttcc tgttctatcc actgtcctag gtttggaatt 120
tcagctgggc agcttcgac ccaccacctg tagccactga ggttcacagg ctctgtttga 180
ggtgttgtgt tacgattgtg ctgggttggt tttgctgctt attgagtgat gaagagcgaa 240
ctatattagc tactgagata ttctgcaga aaacoggaaac gagatgtctt cagctcaaca 300
tgctatcaaa cttgtgaaat aaacagcgtt gttgctcaca caaagcctgt ttggtagtct 360
cttcacatgg acgcgtgtga catttggtgc tgaagaccg ggacaggagg actccttcag 420
gagacgggtc cctgttcctt gccctcactc cgtgaggaga tccacctacg acctcgggtc 480
ctcagaccaa ccagcccaag gaacatctca tgaatttcaa atcggattcc caactatatg 540
aagacacct agctggacga tcagttctta ttaagaacct gacctcaa actctacaac 600
ctcgatggac tggacctac ttagtcatct atagtaccc aactgccgtc cgcctgcagg 660
atcctcccca ctgggttc 678
```

<210> 160  
<211> 932  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:758009.3:2000SEP08

<220>  
<221> unsure  
<222> 880, 883, 892  
<223> a, t, c, g, or other

<400> 160  
gcaaatgagt gacagtctct acttacagac aaagtggagac gtcaggcatt gagacatagc 60  
tccatagaat tcagtttctg agaaccagcc agaagcatgc agtgacattg cacaatctgc 120  
ctctgaagct ggagatacta gctgcagagc tcaggggagc tgctccacat caccgacatg 180  
aaggggaacag gcatcatgga ctgtgcgccc aaggcactcc tggccagggc actttatgac 240  
aactgccctg actgctctga cgagctggct ttcagcagag gggacatcct gaccattctg 300  
gagcaacacg tgccagaaag cgaggggttg tggaagtgtt tgctccatgg gaggcaaggc 360  
ctggccccctg ccaaccgcct ccaaatacctc acggaggtcg ctgcagacag gccgtgcccc 420  
cccattctctg agagggcctgg aagaagctcc tgccagctca gaggagacct atcaggtgcc 480  
cactctaccc gcacctccca ctccaggccc cgtttatgag cagatgagga gttgggcgga 540  
ggggccccag cccctactg cccaagtcta tgaattcccc gacctccca ccagtgccag 600  
aatcatctgt gaaaagactc tcagctttcc aaaacaggcc atcctcacgc tcccagacc 660  
tgtccggggc tcaactgccga ctctgccttc ccaggtgtat gacgtgccta ccagcaccg 720  
gggccccgtg gtcctgaagg agccagagaa gcagcagtta tatgacatac cagccagccc 780  
caagaaggca ggactccatc cccagacag ccaagcaagt gggcagggtg tcccctgat 840  
atcagtgact acttaagaag aggcgggttac agcacattan canatcctca gnaatcggaa 900  
tggtattatg aactccagt gtctccagga aa 932

<210> 161  
<211> 1052  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:331593.1:2000SEP08

<400> 161  
cggatcgagg tgagaaggaa actgcaagag tggggcagag aaccagagtg tcagagcaaa 60  
acctcctcta tctgcacatc ctggggacga accgggcagc cggagagctg cggccggccc 120  
agtcccgctc cgcttttgaa gggtaaaacc caaggcgggg ccttggttct ggcagaagg 180  
acgctatgac cgcagaattc ctctccctgc tttgcctcgg gctgtgtctg ggctacgaag 240  
atgagaaaaa gaatgagaaa ccgcccgaag cctccctcca cgcttgggcc agctcgggtg 300  
ttgaagccga gagcaatgtg accctgaagt gtcaggctca ttcccagaat gtgacatttg 360  
tgctgcgcaa ggtgaacgac tctgggtaca agcaggaaca gagctcggca gaaaacgaag 420  
ctgaattccc cttcacggac ctgaagccta aggatgctgg gaggtacttt tgtgcctaca 480  
agacaacagc ctcccatgag tggtcagaaa gcagtgaaca cttgcagctg gtggtcacag 540  
ataaacacga tgaacttgaa gctccctcaa tgaaaacaga caccagaacc atctttgtcg 600  
ccatcttcag ctgcatctcc atccttctcc tcttctctc agtcttctc atctacagat 660  
gcagccagca cagtgcgctc agagaacgca aaggagagga gggggagtga aggattttct 720  
cgaaccagcc attccaaact tccggagcag gaggtgcgg aggcagattt atccaatag 780  
gaaagggat ctctctcgac ggcagacccc caaggagtga cctatgctga gctaagcacc 840  
agcgccctgt ctgaggcagc ttcagacacc acccaggagc cccagagatc tcatgaatat 900  
gcggcactga aagtgtagca agaagacagc cctggccact aaaggagggg ggatcgtgct 960  
ggccaagggt atcggaatc tggagatgca gatactgtgt ttccttgctc ttcgtccata 1020  
tcaataaaat taagtttctc gtcttaaaaa aa 1052

<210> 162  
<211> 2019



&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1094174.1:2000SEP08

&lt;400&gt; 162

```

ggccactgtg gagagaacag tcacgtagca ggcacaactt ttccggattt aaagaaaaaa 60
aaaaacctgt ctctacgcct ccattcccag gggcgagctg cctctctggc ggcgagctcc 120
ctctctgtca ccaagctccc tgggcgaagc caatcggcgt cgctggggtc cctgttccag 180
aagtccccgc gaaccactg ggactcagat tctccccata cgccgaggat gggccgtcat 240
gggcgccccg aaaccctcgt tcctgctact tctacggggg ccctggcgcc ttgacgcgag 300
accttggggc gggctcccac ttcccatgag gtattttctac accgatcagt gtcccggccc 360
tggccggcgt gggagccctt gtttcatctg cctgtgtggc tacgtaggac gacagcgcat 420
gttcgtgctg gttcgatcag ctgatcgccg cgagccagag tgtatggacg ccggcggggc 480
ccgtgcgata gacgcaggag gggccgctcag gtcacgggga cctgggtaga cacaagatct 540
ccaatgacgc aacgacacat gacttacctg agatgaacct tgttgacact gcgctcctgc 600
ttactacaat ccaggatgct gatggtctgg gtctcacaat catccagagg atgtaacggc 660
tgcagaccgc tggtagccgg acggggggcg gcacattgca ctgcccgagc gggcaattaa 720
atccaagtac gccctacgac gggcaaggat ttactttcgc cctgaaacga ggtacctgga 780
gctcctggac acagccggcg acacggggcg gctcagatgc acccaggcgc aaagtgggat 840
gtgcggcccc gtgaggcgga gcagcggagg agcctaacct ggagggacga gtgcgttgaa 900
agtggctccg cagatacctg gcgaaccggg aaggcaaaag ctgggagcgc gctggaacac 960
ccaaagacac attgtgacct accatccgtt ctctgacctt gaggccacct tgaggtgctg 1020
ggccctgggc ttttaccctt gcggagatca cgttgacctt ggcaacgggg atggggagga 1080
acagaccccc gacacaagag cttgtggaga ccaaggcctt gcacggggat gggaccttcc 1140
agaaatgggc cgctgtggtg gtgccttctg gagagggatc agagagtaca catgccacat 1200
gtgcaagcaa cgaaggggct gcccagccc ctcactctg agagatggga ggcaagtctc 1260
cccaggccgc aacaacgcca tcgtgggcat cgttgcctggc cctgttggtc ctggctgtct 1320
ctagctgtcc taggtatgct gtggtcgctg ctgtgaaagt ggaaggaaga agagctcagg 1380
tggaagaagg gtgaaagata gaaaaggagg gagttaacct caggctgcgt gccaacgacc 1440
aaatgctccc gggctcttga agtggtcgct tccaccagcg ttggaaagcc ctgagaccag 1500
ctgtcttttt tgagggactt gagatagcca ggaattcgtt cacagccgtc cccctttgtg 1560
acattccaag agccctcttg gcgatctctg ttctgtgcca agagcacctt gtacagtgtc 1620
tgcgtccctt ggttatgcat ataattgtga tggagggtgg aggaggacca cgccccaccc 1680
cccgtgtcca ccttgggagc cccctggttc tcccatgtct gacacctgtg tttctcctct 1740
cccagtcac tcttctctgt gtccacagaa gaggtggggc tgtgatagtc tccactatct 1800
ctgtttctca actttaacgt gccactggag gctgccaacc ttctttactt tccccttact 1860
tgaaaaataa gggatcctgg aatataaaat ttgttttctc caaaaaattg cttatgaagc 1920
gttgatggg aattcaatta attaaagccc aattcctaga aaattggaga gagccaatat 1980
aagacctgag aaccttccca gaaaacacaa aaaaaaagg 2019

```

&lt;210&gt; 163

&lt;211&gt; 532

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:814362.1:2000SEP08

&lt;400&gt; 163

```

cagcaggcag gggcagcaag atggtgttgc agaccaggt cttcatttct ctgttgcctc 60
ggtatctctgt cctgactaca ggtgcctacg gggacatcgt gatgaccagc tctccagact 120
ccctgggtgt gctctctggc gagagggcca ccatcaactg caagtccagc cagagtgttt 180
catcacgctc cgacaacagg aattttcttag cttggtacca acagaaacga ggacagcctc 240
ctaagggtgct cattaacgag gcattctaacc gggagtccgg ggtccccgag cgattcagtg 300
gcagcggggtc tgggacagat ttactctca ccatcagcag cctgcaggct gaagatgtgg 360
cagtttatta ctgtcagcaa tatttttagtc ttctctctac tttcggcgga gggaccaagg 420
tggagatcaa acgaactgtg gctgcaacct ctgtcttcat cttcccgcga tctgatgagc 480
agttgaaatc tggaaactgcc tctgttgtgt gcctgctgaa taacttctaa tc 532

```

<210> 164  
 <211> 545  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:219542.1:2000SEP08

<400> 164  
 ggctgtaaaa cggaagggtt cggaatttgc ctctgcgcgcg tctttttttg cctgttacct 60  
 gtgacgtcct tggaaagcaga atctgaaact ttctgaggag agcatttgag ctccagattt 120  
 ctaacagcct ctttgcaaca aaatagacca gtagctgaag gcaactgcaa tccttcata 180  
 tgtttccctt ggccagaaat gcactaagca gtctcaagat tcaaagcatt ctgcaaagca 240  
 tggcaagaca tagccatgta aaacactcac cagattttca tgataaataat ggtaatgctg 300  
 tgctagccag tggaaactgct ttctgtgttg ctacatgggt gtttacagcc actcagattg 360  
 gaatagaatg gaacctatcc cctgttggca gagttacccc aaaagagtg aaacatcagt 420  
 aaccatcaca gttgctgtta tgacagaatt gtttaaaaaa ccaacttgtc atgtaagcac 480  
 tctactgctt attaaaatat agcacaattg aaaaaataaa atgtgtttta aatctttaaa 540  
 aaaaa 545

<210> 165  
 <211> 512  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:726197.1:2000SEP08

<400> 165  
 ccagcccgga atcgagagca atggcgctcg ccaaggccaa ggagatcggt gcctccgctc 60  
 ccctcgctgt cttcagcaaa acttcctgtc cttctgcgt ccgtgtgaag cagttgttcg 120  
 agaagctcgg agcaagctac aaggccattg agttggataa ggaaagtgat ggagctgagc 180  
 tccagaatgc cctgaaggag tggactggac agaggactgt cccaaatgtc ttcataaatg 240  
 ggaagcatat tggcggctgt gatgatacta tggcactgaa caatgatggg aagctgggtg 300  
 ctctgctgac tgaggctgga gccatcgccg gttctgcctc gaagacaacc atcactgctt 360  
 gacatgtgtg ttaccattcc cgtgcttaaa aataatgatg ctttgcaagt gctgaacttg 420  
 cagatgctat gtgggagtga ctgttcgttc cctgtgttaa actccctaga catcagtcga 480  
 gtcgtgttgc tgtggtggtt gtgaaatcta ca 512

<210> 166  
 <211> 1044  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1075314.1:2000SEP08

<400> 166  
 ccggaactca agatggctgc gctcttgctg agacacatcg gccgccattg cctccgagcc 60  
 caccttagtt ctcagctctg tatcagaaat gctgctcctt tgggaaccac agctaaggaa 120  
 gaaatggcac ggttctggaa taagaacacg agttccaacc gtcctgtctc tccccatttg 180  
 actatctaca ggtggtctct tcccatggca atgtctgttt gccaccgagg ctctgggata 240  
 gccatgagtg gaggggtctc tctttttggc ctgtcggcac tgctgcttcc tgggaacttt 300  
 gagtcgtatc tgatgcttgt gaagtccctg tgtttggggc cagcgtgat ccatgcagcc 360  
 aagttcgtgc ttgtctttcc tctcatgtac cactcattga atggggtccg aacttgatg 420  
 tgggacctag ggaaggcct gtcgatttcc cagggtccagt tgtctggagt ggccggcctg 480  
 ggtcctggca gtgctgtcct ctgcaggact ggcagccata tgaagagctg ggattccac 540  
 atcgtcctg tgcacatca cactgatctc tattcctgtt tgtcactcct ccctccagcc 600  
 accaagggtc tctgatttg tttagatgtc atgtgcttcc gatcccttg gggcacagta 660  
 gagaagctta cagaactgta atagtagaaa gaatccgttt tgccctaggc ctaggagccc 720

```

tcattctctt ccacttttga actctgattt gtgctgaggg tcagctttgt gctgcttctt 780
gctgaagaca gtggaaacaa tgccagttct gtggctgccc tgagtgccac tgccgtgtggg 840
ctgcaggctt aaaggacaac ttcatgttca atttggtcagc tcagggcctt taagcaacc 900
atcacagaca gtgaactgag agaagagaaa tggaggtgga ggggatcatc ctgccagat 960
agaggggatg aaagagacag acgtggatct tgggagatat gacattgggg gaaagacaga 1020
cagctttcca tgtatgaagg aaga                                     1044

```

&lt;210&gt; 167

&lt;211&gt; 693

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:437883.1:2000SEP08

&lt;400&gt; 167

```

cggagaaccg cgtaccctgc catgtttctc ttgtgcgctg cgggccaggc gtccggcctc 60
accgctcagt ggggaagaca tgcaaggaat ttgcataaga cagcagtgca aaatgggtgct 120
ggaggagcct tattttgtgca tagagatact cctgagaata acccagataa ctccatttga 180
tttcacacca gaaaactacg agaggtatag aggcaatagt ctagaaacta cccggaaggg 240
cacagagcag cagctgtggc ttccagtgtc ggacctggcc caaaggcaga atgggtggct 300
gcctatctct gccatgaaca aggtggcaga agttttacaa gtgcctccaa tgagagtcta 360
tgaagtggcg actttttaca caatgtacaa tcgaaagcca gttgggatag taccacattc 420
aggtctgcac tactacacct tgcattgttc gagactctga ctagctatcc tggagacctc 480
ttcagagaaa gctgggaata aaagttggag agatatacgc ctgagcaagc ttttctactc 540
tatagaacgt agaattgttt ggggcctgtg taaaacgcac cgatggttca aataaacgac 600
gactactacg aggatctgac acccaaggga tattgaagga gattattgat gaactcagag 660
ctgggaaagt tccagaacca gggccaagga gtg                                     693

```

&lt;210&gt; 168

&lt;211&gt; 1770

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:336265.1:2000SEP08

&lt;400&gt; 168

```

tccgggagaa ccaggagaga aaggagtccc aggcaaggag ggggtccctg ggaaggcctg 60
gagagcctgg attcaaagga gaaaggggag atcctgggat caaagggtgac aaaggacctc 120
ctgggtggaaa aggccagcct ggggaccctg gaatcccagg ccacaaaggc cacacaggcc 180
tgatgggtcc ccaaggacta cctggggaga atggaccagt tggaccccca gggcctccag 240
gccagccggg atttccagga ctgagggggg agtctccatc catggaaacc ctgctcgccg 300
ttattcaaga agagctgggg aagcagcttg aaaccagact cgctacctc ctggcccaga 360
tgcccccgcc gtacatgaag tcatctcaag gcagacctgg gccccaggg cccctggaa 420
aagatgggct tccaggccgg gccggcccca tggggggagc caggctcgcc tgggcagggg 480
ggtctggaag gacctctggt acccataggt cccaaagggt agcgaggagc caaagggtgac 540
ccaggtgcac ctggagttgg cctccgaggg gagatgggac cccctggaat ccaggtcaa 600
cccggggaac ctggctatgc taaagatgga ctctcctgga tccctggccc tcaaggggag 660
acaggaccag ctggacatcc tggcctccca ggacctccc gtccccagg ccaatgtgac 720
ccttcccagt gtgcctactt cgccagcctt gctgcccggc cgggtaatgt gaagggtccc 780
taaaggactc tggaaagcca gaagactgca gtggatttct gaaacttgaa ctgagagccc 840
agtgggaagc cagaggtctt gaaagacttc agccatgtgt tcctttggtt gctgtcttgt 900
cttttatcgt ttgctttttg ttttattttc ttgagagacc tcaaaattat taaatccaac 960
agacgctgcc ggtcggtcag attattatta atattattgt tgttgtaaat tattattatt 1020
atttcatatg ctgatgcttt gtgagttctt ttccactcct ttaaagttgg gaaaacttga 1080
ttcgtggggc aggagattgt ttcttcattc ttctgacagc ccccatctga cgcgtaactg 1140
cccattttaa ggaaactctt ggtgctacaa aaccctgacc agacacttgg caaatttacc 1200
tctttcttca aaagaaaaac tttaagaaaa tgagccaatg ggcttcattc tcagtcagtc 1260
ccggagatca cccaggagaa ataatacaaa caccaccact gtccagagag agtaaagaag 1320
cagaaagaga aagaatttgc aaccatgagg aatgttccca cctcccagc ggacgtgcat 1380

```

```

ttggaaaaca cagaatcagc cctcaggggtg cactccagcc acctcagtg tctaagctca 1440
cagaagtgaa ataattgtctg tgggttggca atggctttgt gggatcatat gtcttggcca 1500
aagatgggaa aacctatggt gaagaggcag cccttgagtg ttaatttgtc ttctaaactg 1560
tgtaaggccc cttcaagttc ctcttgttgg tttcaattat attaatata aaacaagtgg 1620
atgtggtgac catccacttg tgtttcccta atgatgggca gttggccagg gcactgacca 1680
gagctgggaa atttgtatct ccaaggcggc tctgtctctg aaataaatgg catcaagtgc 1740
atgtgtgtat gcgacatgcc ctgcctgaac 1770

```

&lt;210&gt; 169

&lt;211&gt; 1414

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:407788.2:2000SEP08

&lt;400&gt; 169

```

aatatttcog gctattcctt acatacctta catatttaca tatgtataca tgcataattca 60
catccttaac gagcttcatg atgcaattat acctcctgaa atattttaa atgaacttgta 120
aagatttctt ttgcttttag agaggtgaaa agggagaacc tgggtgtccga ggtgccattg 180
gatcaaaagg agaactctgg gtggatggct tgatggggcc cgcaggctct aaggggcaac 240
ctggggatcc aggtcctcag ggacccccag gtttggatgg gaagcccgt tgtattctgc 300
ttctttcaca gttattacat tttcattgat tttgtctcat gttctttatg tgcattgtat 360
aaccctaac actttgatgc aggggaagaga gttttcagaa caatttatc gacaagtttg 420
cacagatgta ataagaggta ggtatcaaat atagcatttt aataactttt ctcaaattac 480
aaaagtaata caaattctta ggacataatt tagaggatat aagtgagggtg aagaagaaaa 540
cagataacat ttataagcct catctacttt tcttcaatga gtatttttat aaagatttat 600
catataact cttttataat ctttttaata aagacattta ttattttgcc attatttgaa 660
ttttcttttg tagcatattt ttgtataatt tagttcataa atatgaatat agtccaccaa 720
tttcttttgg gggcactttc actttttccc ctattttaag caattgctct aataaaaatta 780
ccgcacagag caagtgaata tttccatagt caaattccta gaagtgaat tattggccaa 840
gaaatatgca catattcaat gttcgtagtt ttttctctgt aaactgaaaa aaattaaaaa 900
tataattaac accaactcac catttaaaaca taagggtatgg ggtttctttg tttttgaaag 960
cccagctacc agtcttactt cagagtgga gaattagaaa ttgtgatcat tgcctgtccc 1020
aacatggctc cccgggtatt cctgggccac ctggtccgat aggccagag ggtcccagag 1080
gattacctgg tttgccagga agagatgggt ttcttggtat agtgggtgtc cctggacgtc 1140
caggtgtcag aggtataaaa ggctaccag gaagaaatgg ggaagagggt agccaagggt 1200
ttgggtatcc tggagaacaa ggtcctcctg gtccccaggt tccagagggc cctcctggaa 1260
taagcaaaga aggtcctcca ggagaccag gtctccctgg caaagatgga gaccatggaa 1320
aacctggaat ccaagggcaa acaggcccc caggcatctg cgacccatca ctatgtttaa 1380
gtgtaattgc cagaagagat ccgttcagaa aagg 1414

```

&lt;210&gt; 170

&lt;211&gt; 1430

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1326925.1:2000SEP08

&lt;400&gt; 170

```

caaagaagtg aacctgggag gaaggtgtga ggtagtctca gtggtgggtga tcaaagaagt 60
gaagactgga atactatgtg acatggcctc agtgggtgtg atcgacagaag tcaagctggg 120
agtactgtgt gatcgggact gagtggctgt gattgaagaa gggaagctgg gactactgtg 180
tgagggtgtc tcagtgggtg cgatcaaaga aatgtagctg ggagtactag gtgagatgat 240
agaaggagtg aagcctagaa tactgtgtga catggtctca gtggtgggtga tcgcagaagt 300
caagctggga gtactgtggg atgtggactc agtgatgggt gttaaataag tgaagctggg 360
agtactgtgt gaggtgggtc cggttgtagt gatcgaagaa gtcaagctgg gaatactgtg 420
ggatgtggcc tcgatgggtg tgatcaaaga agtgaagctg ggagtactgt gggatgctgt 480
ctcgggtgta gtgatcaaag aagtgaggct gggagtactg tgtgaggtgg tcttgggtgt 540
ggcaatcgcc aaagtgaacc tgggagtact gtgtaaggtt gtcttgggtg tggcgatcgc 600

```

```
aaaggtgaat ctgggaggac tgggtgaggt ggtcttggag gtggcaattg aagaagagaa 660
gccgggagta ctgtgtgaat gggctctcagt ggtggtgctt gaagaagtga agctgggagg 720
actgtgtgag gtgggtctcgg tgggtggcgat cgaagaagtg aagctgggag tactgtgtga 780
ggtggtcttg gtgctggtga tcaaagcagt gaagctggga ggactgttgg aggtgggtctc 840
ggtggtggtg atggaagaag tgaagctggg agtactgtgt gaggggctct cgggtggctcg 900
gattgaagaa gtgaagctgg gactactgtg tgaggtgggtc ttcttgggtg tgattgaaga 960
agtcaagctg gacgtactgt gtgaggtggt ctgagtggtg atgatcgagt aggtggagggt 1020
gatcgaagaa gtgaagctgg gactagtgcg tgaggggggtc tcggtgatgg tgatcaaaga 1080
agtgaagctg ggagaattct gtgaggttgt ctgagtggtg gtgattgaag aagtgaagct 1140
gggagtactg tgtgaggtga tctcgatagt ggcgattgaa gaagtgaaga tgggagtgtc 1200
gcgtgaggtg atcgctcttg tggatgatcg agaagtgaag cttgaaggat ttgtgtgggt 1260
ggacttggtg gtggcgatca aagaagttaa gctgggagta ccgtgggatg tgggtcttgg 1320
ggcagtgatc aaggaaactga agctgggagt actgtgtgag atggtctctg tgctgctgat 1380
cgatgaaggg aagctgggag taccgtgtga ggtgctcttg gtggtggcga 1430
```

<210> 171

<211> 1184

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:332655.2:2000SEP08

<400> 171

```
cactggcaga gcaaatatga ctcaaaaacc ggctcctcag ggttgtaaca ttagatgata 60
caggcttggg tcgttacaca tgacaccagt gcctttgttt cattgggctg ggctctctgg 120
aaggtgtgct gctgcctgag ctgctggaaa agcactgaca ggtgtttgct agaaaagcac 180
tcctggagct tgccaccagc ttggacttct agggactttc ctctcagcca ggaaggattt 240
tgatattcat cagaaatacc tccagaagat tcaaggagct gtagaggtga agtaagcctg 300
tgaaggacca gcatgggaat cctatactct gagcccatct gccaaagcagc ctatcagaat 360
gactttggac aagtgtggcg gtgggtgaaa gaagacagca gctatgccaa cgttcaagat 420
ggctttaatg gagacacgcc cctgatctgt gcttgacaggc gagggcatgt gagaatcggt 480
tccttccttt taagaagaaa tgtaaatgtc aacctcaaaa accagaaaga gagaacctgc 540
ttgcattatg ctgtgaagaa aaaatttacc ttcatgtatt atctactaat tatcctctta 600
atgcctgttc tgcttattgg gtatttcttc atggtatcaa agacaaagca gaatgaggct 660
cttgtaacgaa tgctacttga tgctggtgtc gaagttaatg ctacagattg ttatggctgt 720
accgcattac attatgcctg tgaatgaaa aaccagtctc ttatccctct gctcttggaa 780
gcccggtcag accccacaat aaagaataag catggtgaga gctcactgga tattgcacgg 840
agattaaaaat ttcccagat tgaattaatg ctaaggaaaag cattgtaatc ctgtgacca 900
caccgatgga gatacagaaa aagttaacga ctggattcta tcttcatttt agacttttgg 960
tctgtgggac atttaacctg gatgccacca ttttatgggg ataatgatgc ttaccatggt 1020
taatgttttg gaagagcttt ttatttatag cattgtttac tcagtcaagt tcaccatggc 1080
cgtaatcctt ctaagggaac cactaaagtt gttgtagtct ccacttcagt cagaaactga 1140
tgtttcagct aggcacagtg gtacatgcct gtaatccag ggg 1184
```

<210> 172

<211> 1101

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1184621.4:2000SEP08

<400> 172

```
aatatttccg gctattcctt acatacctta catatttaca tatgtataca tgcattttca 60
catccttaac gagcttcatg atgcaattat acctcctgaa atattttaa atgaaggttgt 120
aaagatttct tttgctttta gagaggtgaa aaggggagaac ctggtgtccg aggtgccatt 180
ggatcaaaaag gagaatctgg ggtggatgg cttgatgggg cccgccagg cctaaagggg 240
caacctgggg attccaggct ctccaggacc cccagggttg gatgggaagc ccgtatgtat 300
tctgcttctt tcacagttat taccattttt attgattttt gctcatgttc ttatgtgtga 360
tgttataacc cctaacactt tgatgcaggg aagagagttt tcagaacaat ttattcgaca 420
```

```

agtttgcaca gatgtaataa gaggtaggta tcaaatatag cattttaata actttttctca 480
aattacaaaa gtaatacaaaa ttcttaggac atattgtaga ggatataagt gaggtgaaga 540
agaaaacaga taacattttat aagcctcatc tactttttctt caatgagtat ttttataaag 600
atztatcata tatactctttt tataatctttt ttaataaaga catttattat ttgtccatta 660
tttgaattttt cttttgtagc atattttttg ataatttagt tcataaatat gaatatagtc 720
caccaatttc ttttgggggc actttcactt tttccctat ttttaagcaat tgctctaata 780
aaattaccgc acagagcaag tgaatatttc catagtcaaa ttcctagaag tgaattatt 840
ggccaagaaa tattgcacat atttcaatgt tcgtatgttt tttctctgta aactgaaaaa 900
aattaaaaat ataattaaca ccaactcacc atttaaacat aagggatggg gttctgtgtt 960
tttgaagcc cagctaccag tcttacttca gagtgggaaga attagaaatg tgatcatttg 1020
cccgccccaa catggctccc cgggtattcc tgggccacgt ggtccgatag gccagaggg 1080
tcccagagga ttactgggtt g 1101

```

&lt;210&gt; 173

&lt;211&gt; 950

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2051386.1:2000SEP08

&lt;400&gt; 173

```

gaaaaagaca aatttagcat acaaagaaag tagatgctca actgagccag gaggaagtca 60
ctagcaaaaa ggaaaaaac agagctctgg aaaaataatg taaacagtc caagaaaggg 120
accctgtgtc ctgctgaca tctggacagg tatataaaga gcccgggctc agggagctcc 180
acacctgcac ctccctctca cctgctcctc tacctgatcc accctcaatc taccagaatc 240
atgggctgct gtaggctgct ccggtagggc tgtggctccg gctgtggagg ccgtggctcc 300
ggctgtgggg gctgtggctc cggctgtgga ggctgtggct ctggctgtgg gggctgtggc 360
tccagctgct gtgtgcccgt ctgctgctgc aagcccgtgt gctgctgtgt gccagcctgt 420
tctgctcca gctgtggctc ctgtgggggc tccaaggggg actgtggctc ttgtgggggc 480
tccaaagggg gctgtgggtc ctgtgggggc tccaaggggg gctgtgggtc ttatggctgc 540
tcccagtcga gctgtgcaa gccctgctgc tgcctctcag gctgtgggtc atcctgctgc 600
cagtccagct gctgtaagcc ctgctgctgc cagtccagct gctgtaagcc ctactgctgc 660
cagtccagct gctgtaagcc ctgttgctgt tctcaggct gtgggtcctc ctgctggcag 720
tccagttgct gcaatccctg gtgctccag tctagctgct gtgttccgt gtgctggcag 780
tgtaagatct gaggtctga ctgcagactg caggtggcct gactggtgaa gggcccggt 840
gccagcttc cttgcccctg gttctctggt gctccactgt ctccactgtg tctcactgg 900
cttcatccac tccacaccag tgctcccgaa actgactgag gaccccttct 950

```

&lt;210&gt; 174

&lt;211&gt; 673

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:362757.1:2000SEP08

&lt;400&gt; 174

```

cggattcctc ctccctcgac cgtccccgtc cccgtccctc tccggcgcgc aagctcgccc 60
gagatggcaa acgcgagatc ggggtgctgt gtgaatgacg agtgcagct caagttcggc 120
gagctgcagt cgaagaggct gcaccgcttc ctaactttca agatggacga caagttcaag 180
gagatcgttg tggaccagggt cggggatcgc gctaccagct acgaggactt cacaaacagc 240
ctccccgaga atgactgccg atacgcgac tatgatttcg actttgtcac tgcagaagat 300
gtccagaaga gcaggatctt ctatatccta tgggtcccat cctccgccaa ggtgaagagc 360
aagatgcttt atgcaagctc aaacaaaaaa ttcaagagtg ggctcaatgg cattcagggtg 420
gaactgcagg ctactgatgc aagtgaatc agccttgatg agatcaagga tcgggctcgc 480
taggcatcat catgatcatg catcatggac ttggcctact actgtggatt tgtatgccat 540
tatagacttg gtgctgtgaa agactgcttg atgatttgcg ggtttgttgc tgtgtaaaaa 600
aagggtccat ggctcccaga agaccatgaa ggttcggatc tatcatgtaa ttccttgtta 660
tctgccaatt atg 673

```

<210> 175  
<211> 2078  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:406770.1:2000SEP08

<220>  
<221> unsure  
<222> 249  
<223> a, t, c, g, or other

<400> 175  
cccagtcctg gactgccggg cccttcgatg gcagcgtcct gcacggcctc ccaggccgca 60  
gccaaccttc gtgtccgact tcttggagga gttcccgggt gagggtcgtg agtgtgtcaa 120  
ctgcggggcc ctgtccacac cgctgtggcg ccgagacggc accggccact acctgtgcaa 180  
tgcttgccgc ctctaccaca agatgaatgg cgtcaaccgg ccgctcgttc ggctcagaa 240  
gcgcctgtnc tcgtcccgcc gcgcggcct ctgctgcacc aactgccaca cgaccaacac 300  
cacgctgtgg cggcggaact cggaggggga gcccggtgac aatgcctgcg gcctctacat 360  
gaagctgcac ggggtgccgc ggctctggt atgaagaaag aaagcatcca gacacggaag 420  
cggaagccaa agaccatcgc caaggccagg ggctcctcag gatccacaag gaatgcctcg 480  
gcctccccat ctgctgtcgc cagcactgac agctcagcag ccacttcgaa agccaagccc 540  
agcctggcgt cccagtggtg ccctgggccc agcatggccc ccagggcctc tggccaggag 600  
gatgactctc ttgccccggg ccacttggag ttcaagttcg agcctgagga ctttgccttc 660  
ccctccacgg ccccgagccc ccaggctggc ctccgggggg ctctgcgcca agaggcctgg 720  
tgtgcgctgg ccttggccta ggtccccagg ccagcccatt tcagggggaa agcctggaac 780  
agaccacca ctgagtcacc tccgtgcctg ctttgcctca gcacagcaga gaccagcag 840  
ccccccaacc cagagactgg gtctgctgga gtctccacac agtgggtggg aggccttctg 900  
gacagacggc agtcgggccc cagagcaaga aggtgggtga gggaagggct cagcttccca 960  
ccccacgtac agcaagggac tccccagggt cggcccaagg ctccggacca cactggcccc 1020  
ctgcggcgga ggccaacgca gggcaccacc accaccaact tgaattccgt catcaatgct 1080  
caccgtcaat atgtttacaa gttgtagcag ttgggggaaa acagtcaacc tcccagtgt 1140  
aaaccaagat tcccagtgaa gcacctgagg ccaagcaggg gagaggaatg aggggagcag 1200  
ctggacatgg gcctcctgag gcctcggggc tgtccttcat tgcccacatg gatagacgga 1260  
gctgtggtgc agagaacttt tcccgaaca ggtgcaggac tgccagggat cggagtgcgg 1320  
gccgcgcacg gtgccaggat tccgccgagg ggaagccgct cacattgcag tcatcacaga 1380  
cttacgcact tgtttggaca gtttttccag aggggatggg aaagggcctt gttctagctg 1440  
aatctgtgta tcatgacct ttctgacagg cagaatgaat tgtctggtag ccctgtcctg 1500  
accatccaa gcgctgttgg ggctgggtgt gacgtggtca catgtcctgg catatctggg 1560  
gccacgcagt ttagtctctt gtcccaggag aattgttagt gacccctctt tctcttgcaa 1620  
gccccctcca cactgggttg gatgatacct taatgagtga cgctggcgag aggcacccta 1680  
ccgacgcag ctgtgaatgg ccggtgatgt atgtcaggag gccacaggga gcggaggagc 1740  
ggggcaggca gccacagggc cctgcgggga gcacatcctc gcctccgtcc ggctgctgcc 1800  
tttcaacaac aagccctgat ttttccagca atgccagaaa cctggatttt aagtcttcca 1860  
atttgattca aaaatatttt taacattgtg agccagctag acccccagtg caccacccca 1920  
tattgaaaaa cagttgtctg gcacagctt caggagcggg tccggtcatt ctgaaactgt 1980  
ccctccagag gttcttccag ccccaattct atgcgatgtc atcttttcta aaagagacaa 2040  
atgaagccac agggaaagtg aaataaagct tgaacctc 2078

<210> 176  
<211> 804  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1094640.1:2000SEP08

<220>  
<221> unsure  
<222> 158, 727, 751, 772, 774-775, 777

<223> a, t, c, g, or other

<400> 176

```
ccagaggcat ttttccagag gtggagccac ctgaggctat tcctgaaatt ccagaacatc 60
ctccaactga agagttttaa gtcttcaagg aagtcattcc agagggagaa actcctattg 120
ttaaagaag aaaaacacca tcaccaacag caccagangc tatgaaagaa gttgtccctg 180
aaatgaaaat atttgaggat gtacctgaag agccagaaac tccacgtatg aaaacgcccc 240
aagctcctca agaaattatt cctgccaaga cagttccttc caaaaaaaga gaaccccat 300
cagttaaagt gcctgaagct ctccaagaaa ttgtccctga aaagaaaacg cttgtgtgtc 360
ctctcagaaa gcctgaagtc cttcctgatg aagtgcgga ggctctcaga gaagtgtgcc 420
cggaaaagaa agtgcaccc ccccaaagg ctgaagttgt acctgtcaaa gtgcacgaag 480
ctcccaaaga gattatccct gaaaagaaag tgtcggtgtg gcctcctaaa aagcctgaag 540
tcccacctgt taaagtgccga gaagcttcca aagaggttat ccgcaagag aaagtgcct 600
tggctcctcc taaagagcct gaagtcccac ctgttaaagt accagagcct cccaaagaag 660
tagttcctga aaagaaagcg ccagtggctc ctctaaaga gcctgaagtc ccacctgtta 720
aagtgcntga ggctcccaa gaagttgttc ntgaaaagaa agtgccagt cntntncta 780
aaaagcctga agtgccaccc acaa 804
```

<210> 177

<211> 1827

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:001929.1:2000SEP08

<400> 177

```
ccttgacaat ctgtctgtcc gtctgcagct gcgtgactgt ctgtctctgc catgtctctc 60
tcccattgcc gggcccagag gggcttcagc gctcgctcag cctgttctgc tcgctcaagg 120
ggccgcagca ggggaggctt cagcagcagg ggccgcttca gcagcaggag ccttaattcc 180
tttggggggt gcctggaagg ctctcgtggg agtacctggg ggtcaggggg taggctgggg 240
gtgcggtttg gggagtggag ttgtgggcct gggtctctcc tgtgccctcc ggggggcac 300
caagaagtga ccatcaacca gaatccgctg accccactga agattgagat cgatccccag 360
ttccagggtg tgcggacgca ggagaccag gagatcagaa ccctcaacaa ccagtttgc 420
tccttcattg acaaggtgctg gttcctggag cagcagaaca aggtcctgga gacgaagtgg 480
catctgctgc agcaacaggg gttgagtgcc agccagcagg gcctggagcc tgtctttgag 540
gcctgcctgg atcagctcag gaagcagctg gagcagctcc agggagaacg aggggctctg 600
gatgctgagt tgaaggcctg ccgggaccag gaggaggagt ataagtcaa gtatgaggag 660
gaggcccaca ggcgtgccac acttgagaac gactttgtgg tctcaagaa ggatgtggat 720
gggggttttc tgagcaagat ggagttggag ggcaagctgg aggtctctgag agagtacctc 780
tacttcttga agcatctgaa tgaagaagag ctgggcccag tccagacca ggccagcgac 840
acgtctgtgg tgctgtccat ggacaacaac cgctacctgg acttcagcag catcatcact 900
gaggtccgct cccggtacga ggagatcgcc cggagcagca aggtctgagg tgaggccttg 960
taccagacca agtaccagga acttcagggt tctgccagc ttcattggga caggatgcag 1020
gaaacgaaag tccagatctc tcagctacac caagagattc agaggctgca gactcagact 1080
gagaacctca agaagcagaa cgccagcctg caggccgcca tcaactgatg tgagcagcgt 1140
ggggagctgg ccctcaagga cgctcaggcc aaggtggacg agctggaggc tgcctctgag 1200
atggccaagc agaacctggc ccggctgctg tgcgagtacc aggtgctgac gagcacgaag 1260
ctttccctgg atgtggagat tgccacttac cgcaggctgc tggaggcgga ggagtgcagg 1320
atgtctgggg agtgccaccg ccaggtcact atctcctcgg tgggaggcag cgctgtcatg 1380
tctggaggag ttggtggagg cttggggagc acttgtggac tgggtagtgg gaaaggcagc 1440
cctgggtcct gctgcaccag cattgtgact ggaggctcca acatcattct gggctctggg 1500
aaggaccctg ttttggattc ctgctctgtg tctggctcca gcgtggctc cagctgccac 1560
accatcctga agaagacagt tgagtgcagt ctgaagacat ccacaccta ctgagcgacc 1620
cagcagccac ctcttctctg aacacatttg gcccactccc cccatcagcc ggctctgcaa 1680
ggccaactcc gtgtccgctg cccacagccc aagccagccc acagcggatg ctgcaaaaat 1740
caataaagtc tcccctcctg ctgttctgaa tgctctaagt gcttgcacac ctcaccagc 1800
aaaacaaaag ctgtgtgact cccagc 1827
```

<210> 178

<211> 817

<212> DNA



&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:401322.1:2000SEP08

&lt;400&gt; 178

gcgacccaat	gggagtcgcc	gccaggagac	acaaccggcg	gctctccggg	gatttataag	60
gtcccgccag	caaccgcgcg	ttgggattcg	tgggtgtttgc	tggagaccgc	tattgtgtct	120
ggagtcaggg	aaccttccaa	cttcaactac	tggatcaacc	atgtgtgagt	gcatctccat	180
ccatgtgggc	caggccggag	tgcagatggg	caatgcctgc	tgggagctgt	actgcctgga	240
gcacgacatc	cagcccagtg	gcaccatgcc	cagccacaag	gccctgggga	gcaagtgata	300
actccttcaa	caccttcttc	agggagaccc	agcctggcag	gcatgtgtcc	tgggctgtct	360
gtggacctgg	agcctgctgt	cataggttgg	catcaactac	ccagtcccc	acagtgggtc	420
ccggggcggtg	ctgtagccaa	ggtgcagcgg	gcagtctgcg	tgctaaacaa	taccacagcc	480
atcactgagg	cctggggccc	gcctcaacca	aaagtttgac	ctgatgtatg	ccaagcgggc	540
atttatgcac	tgttatgtgg	aaccggggca	tggaggaagg	tgaagctctt	ctgggcctga	600
gaaggagcct	gggctgcctt	cggagaagga	ttattgaagt	atggcaccac	gaaatctttg	660
ggacagcggg	ggatgagcac	tgaggagttt	taggggtcct	ggcaccctgt	acaccctggg	720
ctctcgacac	accacgttgt	gtgtgatctc	acattcccct	gtttcctgac	cccctaccaa	780
taccatggat	attgccagcg	gaaaaaaaaa	aaaaagg			817

&lt;210&gt; 179

&lt;211&gt; 2510

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:208748.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 1850, 2149

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 179

caaaaaatac	gtgatcagct	tccgtcttgg	atagatcagg	aacgaagctg	ggccgtggca	60
acattaaaga	ttgctctccc	cacgtcttta	tacagacacc	tctgctttga	agatgcagct	120
ctgtggcgta	cttattataa	taattcaatg	tgtgagcaag	agtttccatc	tatccttgca	180
aagaaagtgt	ccttattttc	gcagattcct	gtagtacagg	tgctaagacc	ggacagattg	240
caaagtgcc	tggctctttt	tgcattgtaa	actctgggac	tgaagaggt	gtccccactg	300
cctctaaatc	tcaaacgttt	atacaaagag	acactggaaa	ttgaacccat	cttgataatt	360
atttctccgg	gtgctgatcc	ttctcaggaa	cttcaagaac	tagctaattg	tgaagaagag	420
cggagagtgt	tatcaccagg	ttgccatggg	tcaaggtcaa	gctgatttag	caattcaaat	480
gctaaaagaa	tgtgcccgc	atggagactg	gctctgtttg	aagaacttac	atcttgtggg	540
atcctggctg	ccagttcttg	aaaaggaatt	gaatactctt	caacctaaag	atacctctcg	600
tctttggctc	actgcagaag	ttcatcccaa	ctttactcct	atgtttacta	cagtcaagtc	660
tgaagataac	atatgagtca	cctccagggt	taaagaagaa	tttaatgcgt	acttatgagt	720
cttggaactcc	tgagcaaatt	agcaaaaaag	ataatacaca	tcgagctcat	gtcactatc	780
agtcttgc	ggtttcatgc	tgcattgtcaa	gaaagaagaa	actatattcc	tcagggttgg	840
acaaagtgtt	atgaattttc	tttatcagat	cttcgggctg	ggtacaacat	tattgacaga	900
ctttttgatg	gtgccaaaga	tgtacaatgg	gaatttgtac	atgggtttact	tgaaaatgct	960
atztatggag	gacgtataga	caactatttt	gaccttagag	ttcttcagtc	atacctgaag	1020
cagtttttta	attcttcagt	tattgatgta	ttcaaccaa	ggaacaagaa	aagcattttt	1080
ccatattccg	tatctctacc	acaatcctgc	agcattttgt	taggactatc	gtgctgtcat	1140
tgagaaaatt	ccagaggacg	acaaacctag	tttctttggg	ctgcctgcc	atatcgctcg	1200
ctcatctcag	cgcattgatca	gttctcaggt	tatttcacag	ttgaggattt	tggccagatc	1260
cataacagct	ggttccaaat	ttgatagaga	aatctgggtc	aatgaacttt	ctcctgtcct	1320
caatctcttg	aagaaactaa	accagaattc	aaaccttaata	catcagaaag	tgccctctcc	1380
taacgatcga	caaggatctc	caatactgtc	attcatcatt	cttgaacaat	ttaatgctat	1440
tcgttttagta	caaagtgtcc	accagtctct	tggctgtctc	cagccaaagt	catcagagga	1500
actactttac	tgagttcaga	agtacaaaaa	ttggcaagtg	ctttatttaa	ccaaaagtgt	1560

```

cctctcgcat ggcagagcaa gtgggaaggg ccagaagatc ccttacaata cctgagaggt 1620
cttggtgccc gtgcccttgc aatacagaac tgggtagata aagctgaaaa acaggctctt 1680
ctctctgaaa acacttgaaac ctatcagaaa cttttccatc caggacacat ttcttaatgc 1740
tcttcgccag gaaactgcaa gggcagtggt tctgtctgtg gataagcctt aaattttagt 1800
cctcatggaa aggtcgactg caagaagcaa agctacaaat taagatgcan tggcttggtt 1860
ctagaaggat gtatgttttga tggaaatcaa ctttctgaaa atgcagcttg attctccag 1920
cgtgtcatca gtgctccctt gttttatggg ctggattcca caggatgcat tgtggtccat 1980
attctccgga tggagtgcac ctctttgcct gtttacacaa gtgctgaaaa gggatcggtg 2040
ggttcccaat attgatgttc catgtggggg caacccaaga ccagtggatt cagtgtggag 2100
cagctctatt cctaaaaaat cagtagaatc taatgacaac aagaagccna atcttcacaa 2160
aagggaacat tgattcttta agctttaaat caaacatgtg gtcagtctaa atttgacaat 2220
gttagttcaa gatattaaca tatagttatg ttgttgatgt cactgaaatt ttaatgtgta 2280
aaagcagcac tgtgcatctt ttaaagttaa taaaattaat gggagtattt gttaaaacag 2340
agtattcttt tgacaacatt aaatatttca tgtgagaaag ttcacttttc cagtgggtca 2400
aaaatttgtt ttaggtcaga gattttccac tggatatatta accaataata aatatttagg 2460
ctgtcatttg tgcataatt atttaataaa agagcattca taattttaaa 2510

```

&lt;210&gt; 180

&lt;211&gt; 2043

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:407242.1:2000SEP08

&lt;400&gt; 180

```

ggggagccgc cgggacacca aaataggagc tgccttggtg gtggagcggc actagctggc 60
ggcttcogag cgcctcttcc aaagatgggc agagggggcg gaggcgtccc cgctcccgtc 120
cgctactagc ccgcggggcca gcgcgcgcgc ccgagccccc gcgggagcca tggctctaaa 180
aggacaagaa gattatattt atcttttcaa ggattcaaca catccagtgg attttctgga 240
tgcattcaga acattttact tggatggatt atttactgat attactcttc agtgtccttc 300
aggcataatt ttccattgtc accgagccgt tttagctgct tgcagcaatt attttaaggc 360
aatgttcaca gctgacatga aagaaaaatt taaaaataaa ataaaaactc ctggcatcca 420
ccatgatatt ctggaaggcc ttgtaaaatta tgcatacact tcccaaattg aaataactaa 480
aagaaatggt caaagcctgc ttgaggcagc ggatctgcta cagttccttt cagtaaagaa 540
ggcttgtagc cgttttttgg taaggcactt ggatattgat aattgtattg gaatgcactc 600
ctttgcagaa tttcatgtgt gtccagaact agagaaggaa tctcgaagaa ttctatgttc 660
aaagttaagc gaagtgtggc aacaagaaga atttctggaa atcagccttg aaaagtttct 720
ctttatcttg tccagaaaga atctcagtgt ttggaaagaa gaagctatca tagagccagt 780
tattaagtgg actgctcatg atgtagaaaa tcgaattgaa tgcctctata atctactgag 840
ctatatcaac attgatatag atccagtgtc cttaaaaaca gccttaggcc ttcaaagaag 900
ctgcctgctc accgaaaata agatccgctc cctaataatac aatgccttga atcccatgca 960
taaagagatt tccagaggtt ccacagccac aatgtatata attggaggct attactggca 1020
tcctttatca gaggttcaca tatgggatcc tttgacaaat gtttggattc agggagcaga 1080
aataccagat tataccaggg agagctatgg tgttacatgt ttaggacca acatttatgt 1140
aactgggggc tacaggacgg ataacataga agctcttgac acagtgtgga tctataacag 1200
tgaaagtgat gaatggacag aaggtttgcc aatgctcaat gccagggtatt accactgtgc 1260
agtcaccttg ggtggctgtg tctatgcttt aggtggttac agaaaagggg ctccagcaga 1320
agaggctgag ttctatgac ctttaaaaga gaaatggatt cctattgcaa acatgattaa 1380
agggtgtggg aatgctactg cctgtgtctt acatgatgtt atctacgtca ttggtggcca 1440
ctgtggctac agaggaagct gcacctatga caaagttcag agctacaatt ccgatatcaa 1500
cgaatggagc ctcatcacct ccagtccaca tccagaatat ggattgtgct cagttccgtt 1560
tgaaaataag ctctatctag tcggtggaca aactacaatc acagaatgct atgaccctga 1620
acaaaaatga tggagagaga tagctcccat gatggaaagg aggatggagt gcggtgccc 1680
catcatgaat ggtgtatatt atgtcactgg aggatactcc tactcaaagg gaacgtatct 1740
tcagagcatt gagaaatatg atccagatct taataagtgg gaaatagtgg gtaatcttcc 1800
cagtgccatg cgtctcatg ggtgtgtttg tgtgtataat gtctaattga atctgcagaa 1860
atgaccaagc aatcactttt ttggagtata gttttataaa aaaagaatgc agggtttgaa 1920
gttccttacc tgataattgt gtctggcaca tgaatgggga tcagtaaatg gtaattccta 1980
accctactgt actcccaaac atgggtgatt atgggtcaaga aaaatcttat atatatacat 2040
ata

```

<210> 181  
<211> 2288  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:403409.1:2000SEP08

<400> 181  
gtggccgagg cagacggaag ccgaacgagt tcctcgggcg ctgcaggatg ggggactcca 60  
aagtgaaggt ggcggtgagg atacgaccca tgaaccggcg agagactgac ttgcatacca 120  
aatgtgtggt ggtatgtggt gcaacaagg ttattcttaa tcctgtaaat acgaatcttt 180  
ccaaaggaga tgcccggggc cagccgaagg tgtttgctta tgatcattgt ttctgggtcta 240  
tggtatgaatc tgtcaaagaa aagtatgcag gtcaagatat tgttttcaag tgccttggag 300  
agaatatcct gcagaatgct tttgatggct acaatgcatg tatctttgcc tatggacaga 360  
ctggctctgg aaaatcttat accatgatgg gcacagctga ccaacctgga ttaatcccaa 420  
gactttgcag tggactcttt gaacgaactc agaaagagga aaatgaagaa cagagtttta 480  
aagtgaaggt gtcctacatg gaaatttata atgaaaaagt tcgagacctt ctgtatccca 540  
aagggaagcg tcagacgttg aaagtacagag agcatagtgt gttgggacct tatgtcgacg 600  
gactttctaa actggctgtc acaagctaca aggatattga gtcgttgatg tctgagggta 660  
acaaatctcg cacagttgct gcaaccaaca tgaacgagga gtagtagcca tcccatgcag 720  
ttttcaaaat caccctcaca catactctct acgatgtgaa gtctgggaca tctggagaga 780  
aagtgggcaa actcagcctg gtgatttag ctggcagtga acgagcaacg aagacaggcg 840  
ctgcagggga caggctgaag gaagggagca acatttaaca gtccctcaca accctcgatc 900  
tggttatctc agctcttgca gatcagagtg ctggcaaaaa caagaataaa tttgttccat 960  
atcgtgactc agttctcact tggctgctca aagacagcct cgggggtaac agcaagaccg 1020  
ccatggtggc tactgtgagt cctgcagctg ataactatga tgaaccctc tcaactctgc 1080  
ggtatgcaga tcgagccaag cacattgtaa accacgctgt ggtgaatgag gaccctaattg 1140  
cccgaattat ccgggatctc cgggaagaag ttgagaaact ccgggagcag ctgaccaaaag 1200  
cagaggcaat gaaatctcca gagctaaagg accggctgga agaactctgag aagctaattc 1260  
aggaaatgac tgtgacctgg gaggagaaat taaggaaaac ggaggagatt gcacaggaac 1320  
gacagaaaca gcttgagagt cttggaatat ctcttcagtc ttcgggaatc aaagtggggg 1380  
atgataaatg cttccttggt aatctgaatg ctgaccagc tctgaatgag cttctggtgt 1440  
actatttaaa ggaacatata ttgatagggt cagcaaatc ccaagatata caactgtgag 1500  
gcatgggaat tcttctgaa cactgtatta tagacatcac gtcagaaggc caggttatgc 1560  
tgactcctca gaagaacacc agaactttg taaatgggtc atctgtctcc agtccaatac 1620  
agctacacca tggggacagg atattatggg gaaacaatca tttcttcaga ctcaatttgc 1680  
ctaaaaagaa aaagaagca gaacgagagg atgaggacca ggtccctcc atgaagaacg 1740  
agaatagttc tgagcagctg gatgtagacg gagactcctc cagcgagggtg tccagtgaag 1800  
ttaactttta ttacgaatac gcacagatgg aggtcaccat gaaggccctg ggcagcaatg 1860  
atccgatgca gtccatatta aacagcctag aacaacagca tgaagaagaa aaacgatctg 1920  
cactggagcg ccagaggctt atgtatgagc acgaattgga gcagctccgg agaaggctgt 1980  
ctcctgagaa gcagaactgc cggagcatgg acaggttttc tttccactcg ccagcgctc 2040  
agcaacgctt aagacagtg gctgaggaga gagaagcaac gttgaataac agcctgatga 2100  
ggctgagggg acaaattggt aaggccaatc tattggtgag agaagctaata tacattgctg 2160  
aggagctgga taaaagaaca gaatacaaa ttaccctaca gattccagcc tccagcctgg 2220  
atgccaacag gaagcgaggc tctctcttta gtgagcctgc aatccagggtg aaaacaaaaa 2280  
aaaaaagg 2288

<210> 182  
<211> 1791  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:450798.1:2000SEP08

<400> 182  
caccaccttc ctctcctcgc tctccctatt ctggacggcc gactcgcttc tgtgagctcc 60  
cctccacctc gcagctaccg gccttagatc cagccatgag ggagattctc cacatccagg 120  
gcggccagtg cggttaaccag atcggttcca agttctggga ggtcgtctgc gacgagcatg 180

```

gtatcgatcc caccggacgc tatgtcggca cctccgacct ccagctcgag cgcgtcaatg 240
tctactacaa cgaggcctca tgcggcccgg cttcgtgccg cgcgtgtgct tcatggacct 300
cgagcctggc accatggaca gcgtccgcac cggaccgtac ggccagatct tccgccccga 360
caactttgtc ttcggacagt ctggccgcgg ggaacaactg ggcccaaggg ccactacaca 420
gagggcgag agctcatttg actccgtgct tgacgttgct agaaaggagg ctgagaactg 480
cgactgcctt ccaaggcttc caagtctgcc atgtcccttg gaggaggcac tggatctgga 540
atgggcacgc attctcattg tcaaagatca gagaggatgt acccctgacg ggtatgactg 600
ctgacagttg tgtctgtctt gtccgctcgc ccaagggtgc agacacctgt tgcttgagcc 660
ataccaatgc aacccctgtc tgtccatcag gttgggtgag ttacgctgat gtaatgcatg 720
gttcttggac atacgatagc acttttatga catttgcttc cgaactctga agctgacgac 780
cccctagtca tttggagact tgaaccactt ggatcagtg c aacgatgagt ggtgtcactt 840
ggtcgcctcc gccttcccag gacagactgg gaactctgag cctggcgcat atgcttgcac 900
gtcaactctg atcccgttcc ctccgctcca cttcttacat ggttggcttt gcgccactca 960
ctcgcgtggc tccgcagcag taccgggtgc ctcactgggt cctgagctaa cccagcagat 1020
gtgggtgatg caagaacatg atgtgtgctg ctgaccacg ccatggctcg ttatctcaca 1080
gccttctgcc atgttcccgt ggcaagatga gcaactaagg ggttgatgag caacatgatt 1140
aatgtocaga acatagaact ctgtcctatt tgtgtggagt ggatcccaa caacgtgcaa 1200
gtccagtgtt gtgtgataat cctccggcgg gggcctctgc catggccgcc accggtcagt 1260
tggccaactg gcacggggca tcccgagaga gtttcccccc cgttgtgagc gaagcagatg 1320
tcaactgctaa tgttcacgga ggaaaggctt tcttgacacat ggtaacacac tgggtgaagg 1380
cacatggaat gagaatggag aggttcaccc gaaagggctg agaaggcaca acacatgaga 1440
cgagatctgg gtgtccgagg gaaccagcag taaccacgga attgccacct gctgaatgaa 1500
ggagggcgaa atacgaggac gaaggaggcc catctccaga cgaggtaaag ggtggtttgg 1560
ggcgctgccc accaagtgtg agccccctgt gtgtccttgt caattccggg ggttgccta 1620
atggtttgcc ctgttgaat tcgctcacat gttattgcgg taatgtaagt agtacattt 1680
cgtgcgtaag caacatgcct gtaatgctag ccacaccaag actttgttca cctggcactg 1740
ctttggatat gccagcccc tggcattatg tgacgagtat ttcattattt t 1791

```

&lt;210&gt; 183

&lt;211&gt; 1529

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:410317.1:2000SEP08

&lt;400&gt; 183

```

gacccatcag ggcttctgta aggetgagtg ggtcccatgc ctgaaaggca gagccaggct 60
gagcctggcc tgagtctccc atgtgtaaca ggtaccagga tgaaatcaac aagcacactg 120
ctgcagagaa cgagctttgtg gtgtctcaaga aggatgtgga tgcagcatac atggggccga 180
tggatctgca tggcaaatg ggcacctga cccaggaga ttgacttcct gcagcaactc 240
tatgaaatgg agctgagcca agtgcagacc cacgtgtcta acaccaatgt ggtgctgtcc 300
atggacaaca accgcaacct ggacctggac agcatcatcg ccgagggtcaa ggcccagtat 360
gagctgattg cccagaggag cgggctgag gccaggcct ggtaccagac caagtatgag 420
gagctgcagg tgactgctgg gaagcatggg gacaaccatg cgggacacca agaacgagat 480
ttgctgagct caccgcgaac tatccagagg ctgcaggggg aggetgatgc agccaagaag 540
cagtgtcagc agctgcagac ggccattgag gaacgcggag cagcgtgggg agctggcact 600
caaggatgct cagaagaagc tcggggatct ggatgtggcc ctgcaccagg ccaaggagga 660
cctgacacgg ctgctgcgtg actaccagga gctgatgaat gtcaagctgg ccctggacgt 720
ggagattgcc acctaccgca agcttctgga gagcaggag agcaggatgt ctggagaatg 780
tccccagtg agtcagcatt tctgtgactg gcaactccac cactgtgtgc ggagggtggc 840
cagccagctt tggaggtggc atctccctgg gtgggagtg gggggccacc aagggtggat 900
tcagcacaaa tgtgggctat agcaccgtca agggaggggc agtctcttgc gggcacctcc 960
atctctgcga agaccactac ggtcaagacg tccagcccag aggtattage tgctgagccc 1020
tgcaaggggc ccttgaatc ctgtccctgc cctcctcacc ccacctcgtg ctgtcctttc 1080
cagtcacttc tcaggagcag gaacagccag gggaccctca gaaccagggt gtattttcat 1140
taccagactt atttgcatt tgggaagcgc tcaaatctac tcagggtttc tccttgggac 1200
tgcggtgagg aggggggggg gaggggtatg tgggtccgcg atgtcgttgt agtgatgggt 1260
cttgggtgag cttgggggtg accttttgac cccagagagg agtgctgaat ttctcaagcc 1320
attaggagag agagaaattg ggagtgttcc ccaaagacc ttcaacctcc ccagtcctcc 1380
accagaccga cctctcccc tgaatctacc cacatcccc ttcccgtgtc tgtgtctcac 1440
taaagtgggt caacgtgcag gcatagctgt ctgttggctg cagtcgtgtc ctgcctggga 1500

```

aaaacaacaa aacggggggg ccgcacaaa

1529

&lt;210&gt; 184

&lt;211&gt; 1308

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:340268.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 100

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 184

```

ccccgactat tcttgtatca gaaatgtaac tcacccactc accaccagcc cacactctgg 60
gttctctctat cactccactg cctgctctcg gctgatcacn atggaagaag agatcatggc 120
acttgtcatt gacaatgggc tccagcacat gcaaagctgg ctttggctggg tatagttccc 180
cgagccatgt tccccctccat cactgggggt tattccctgg gcaccagggc atgatgggtg 240
gcataggcca gaaggactcc tacctgggtg atgaggccca gagcaagtgc agcatcctga 300
ccctgaaagt gccccgtgga ccaccgcac atcaccaact gggatgacat ggagaagatc 360
tggcaccaca ccttctacga ttagctgcac gtggctccca aggagcacc agtgctgcat 420
gactgaggcc ccgctgaacc ccagggccaa cagaaagaag atgactcgga ttgagacctt 480
caacacccca gcaaggtaca cagccatcca ggctgtgctg tccctgtaca cctctgggta 540
caccactggg attatcatat acagctgaga cccacacggg gcccgctctac aagggtctatg 600
ccttccccca tgccatcctg tgtctgggat actggcacct gaccaaccac cccatggaga 660
tcttcaggga gccagctac tgcttcacca ccacagccaa acgggagatc ttgtgcaaca 720
tcaaggagaa gctgcggaaa tggccaccac agcgtcctcc tcctccctgg agaagatcta 780
ggagctgtcc acagtgggtc cagtgtctgg aggtgctggt ccaggcctcc ttctgggta 840
tggaatcttg ctggcatcca cgagaccacc ttcaactcca tcatgaatgt tgtgactgtg 900
gacatcctgc aaagacctgt acgccaacac ggtatgctgtc gtggcgccac caccatgtac 960
cctgggcatt gccgactgga tgcagaagat cagcacctgt gtgcccagct ccaagaagat 1020
catcaccttc cctgagcgca agtccttgtg tggagttggc agcaccactc ctggccgcct 1080
tccagcagaa tgtggagtca tcaagcacgg agtaggagtg aaatcaagcc cctccatcac 1140
ccacctgcac aaatgctttt aacaatggag ctgcaagca gaacaacgtg tcaatgtgct 1200
tcaatggagt taaatgcaga agtataaaat tgccccctggc aaaagcaaac acctcatgct 1260
agcctccaat aaaccagaat aagcccttcc caaaaaaaaa aaaaaagg 1308

```

&lt;210&gt; 185

&lt;211&gt; 1066

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2051671.1:2000SEP08

&lt;400&gt; 185

```

taacccaaag tatcatggat gtgggctggc cagagctcca cgcaccgccc ctggataaga 60
tgtgtaccat atgcaaggcg caggagtctt ggctgaacag caacctccag catgtggctg 120
tcattcactg caggggcggg aaaggacgca taggagtggg catatcatcc tacatgcatt 180
tcaccaacgt ctcagccagc gccgaccagg cccttgacag gtttgcaatg aagaagtttt 240
atgatgacaa agtttcagct ttaatgcagc cttcccaaaa aacggtatgt tcagttcctc 300
agtgggctcc tgtccggtac ggtgaaaatg aatgcctctc cctgttctct gcattttgtc 360
atcctcctgg gccgctctgc cgggctgggg ctgagcagcg atcctgcttt gtcccagaag 420
tccagaggga tcagccccag aacacaccct cctccccggg acgcacgcag ctttctggag 480
gctgaggaag gcatgaagaa gtgggctcca cctgctggcc gactgagaaa agaattttcca 540
gaactcggtc ctattttaca gattgagaaa ctatggttca agaagagagg acggggcttg 600
agggaatctc ctgattctcc ttatatgacc tcagaactga cctatactaa tactagtgtg 660
gaagggtctt ttacgcgctc taggaatgtc agggctcatc cagtccccgt gatgccatga 720
cttgtgagca gtcagctgat ggtatgatat gcacggtttt cctccatcca ggtttgacat 780

```

```

acaggggggat gatctgggtc ccattctggt cttaagaccc caaaacaagg gtttttttcag 840
catccaggat ctggagcctc tatctggatt agtgtcgtaa cctctgatgt gcctcccgtg 900
taccocatct gtccagttag ctccagcccc atccaacctc acaggggtggc cacagggtatt 960
actgaggggtt aagacccttta gaacgtgggt ctaccacccg gataagagct caataaagtg 1020
ttgttccttg tccacatcaa agaaacaata aaagggggggg ccccca 1066

```

&lt;210&gt; 186

&lt;211&gt; 294

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:998844.1:2000SEP08

&lt;400&gt; 186

```

gggccgagga gcctgggccc ggccggggcg ggactactcc ggagtcagga ggcagcagcg 60
gcggaggacg aggatctctg tccagtcagc gccgctcgga cgcgcggcgc accatgggct 120
gctgcaccgg acgctgctcg ctcatctgcc tctgcgcgct gcagttgggc tcagcattag 180
agaggcagat ctttgacttc cttggtttcc agtggggcgc tattcttgga aattttctac 240
acataatagt tgtcatattg ggtttgtttg ggaccattca gcacatcccc cttt 294

```

&lt;210&gt; 187

&lt;211&gt; 732

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1043787.1:2000SEP08

&lt;400&gt; 187

```

ggcgtcttga gccgttatt gaagccagca ctgctgtgcg gggctctggc tgtgcctggc 60
ctggctggca ccatgtgtgc atcccgcgat gattggcgct gtgcgcgctc catgcacgaa 120
ttcgcagcca aggacatcga tgggcacatg gtttgcttgg ataagtacag gggttgcgtg 180
tgcatcgta ccaacgtggc ctgcgaatga ggcaaaaccg acgtaaacta cactcagcta 240
gtcgatctgc atgcccgata cgccgagtgt ggtttacgaa tcctggcctt cccttgcaac 300
cagttcggga ggcaggagcc aggaagtaat caagaaatca aggagtttgc agccggctac 360
aatgtcaggt ttgacatgta cagcaagatc tgtgtaaatg gggacgatgc ccacccactg 420
tggaatgga tgaaagtcca gcccaagggc aggggcatgc tgggaaatgc catcaaatgg 480
aacctttacca agtttctcat tgataagaac ggctgcgtgg tgaagcgcta tgggtcccatg 540
gaggagcccc aggtgataga gaaggacctg ccgtgctatc tctagcccta caagtgtgtg 600
ccctctgacc gagccccctt gtcctgtgac cctctggagc cttccacccc ggcactcatg 660
acggtctgcc tgaaaaccag cccgtgggtg gggcagtcgc gaggacctgg cgtgcatccc 720
cgccggagga ag 732

```

&lt;210&gt; 188

&lt;211&gt; 582

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1098931.16:2000SEP08

&lt;400&gt; 188

```

aatgtgtcat tgacatgtta ctaacatggt attaacatcc atccctacga tgttcacatt 60
aagaaaaaaa aaacaggggc tttcagtga gggaaactga ttcgtagcct cactaaagta 120
ttttttaagg gatttaaata cctgtacctg actccccaag actacaccag aatcagctcc 180
ctgaactccg tccactgtaa acacatcgag gaaggaggag agtccaggta aataacttat 240
caggtagctc cttaattttg gctcgtgggt aatttcagct gtcttcttcc ctccggggctt 300
gtggcgatat tttccagtgg ggtggtgggt tttgtatgat gggaaccaat gcctgttggg 360
ggcactgatg ccttaaacat ttgttaaact ctccaagggt agtccatcct tgtcattcta 420

```

aaatgagtga ctcctttcac aaaaaatagc acgtcatgta gggagaatga gattggattc 480  
tcgtgaatct cctaggagcc gcaggctcac attttgtgac caaactgatg cagcgggtgac 540  
ttggcaaaact gggacggctc ggggtggggcg agagaggtct gt 582

<210> 189

<211> 257

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:199423.2:2000SEP08

<400> 189

ctgagagcag gacgcatagt gagttgtctg ggacttggct gcagctggca gtacttcac 60  
ctcctcatca tcacggacgg ggtcatcagt gacatggagg agacacggca tgccgtgggt 120  
caggctttcca agctgcccac gtccatcatc atcgtgggcg tgggcaatgc ggacttcgct 180  
gccatggagt tcctggatgg ggacagccgc atgctgcgct cccacacggg ggaggaggca 240  
gcccgcgata ttgtgca 257

<210> 190

<211> 810

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1075297.1:2000SEP08

<400> 190

gcgcttcctc ctgggattca gtcatttaaa gattgagacc aagattgaaa gcatggctga 60  
cctcaagcag ctcatggaca acgaggtggt gatggccttt acctcctatg caacgatcat 120  
tcttgccaag atgatgttcc tgagctccgc gactgcattc cagaggctaa ccaacaagg 180  
ttttgccaac ccggaagact gtgctggcct cggcaagggg gagaatgcca agaagttcct 240  
tcggactgac gagaagggtg aacgcgtgcg aagagccac ctgaatgacc ttgaaaacat 300  
cgctcccttt ctccgtatcg gcctcctgta ctccctgagc ggaccggatc tctctacagc 360  
cctcattcac ttcagaatct ttgtgggccc tcggatctac cacaccattg cttacttgac 420  
tccccttcct cagccaaaca ggggcttggc attttttgtt ggctacggag ttactttgtc 480  
aatggcttac aggcgtgctc ggagcagact gtacttgtaa agagatttgt gaccgtcacc 540  
ctctgattga tttaaaaaaa taaaggattc tatatttcta gtgtattaaa aaactttcta 600  
aggttttatg tatgaaagga gcagagaaat caggaatcag ggaaaattct gtttgaagac 660  
taccatccac aggcctctacc cttttgtacc tgggttagac atttaacatg accggcctta 720  
gctatgctgt ctaactccta aagtactttg tcctaagtct tgagtgcctt ataaattgta 780  
acattccttc tacatttcct gtggtgcac 810

<210> 191

<211> 780

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1043321.1:2000SEP08

<400> 191

ggatgaagca catgagcagg aggcggggct tgtaataacct cctctggaac atcaccagct 60  
tctcagcttt caggctcagac atgttcagct ttccggcctt tctctttgac agccgggtgt 120  
ttgcgcacaa gcagccaacc cacgtgagag aagaaaaagc cacggcgga gttgtgagg 180  
tcggcggtgt tctctgagaa cttgtgggtg gcgcggtgat cccgggccc ttcgtacaca 240  
tcgttctgga atgccatggt gttggcgatg atgagaaaga gcttacaggg cgcagctcga 300  
gccttgtaag tcctgtggct ccacaggcga tgagcccggt ctgtgatgcc cagggtctctg 360  
attacgtagt acaaatatgc aaagaggcag gtgtagacct tgcaggagg gaccagtggt 420  
atcccgtaca gggctccaat gtgcagcagg gccatgagga cgatgttctc tccagacgta 480

```

ctccagcttg ggcggggggcc cctcctcctc ctggtagctg gggtcgtata ggtcatcttt 540
aatttcaggg cgaacatctg ctccccagtg gtgaggattc ttttcaaact tctctcctcc 600
attctgctgt ccccccagaag gtggcgctgt gattgtggtg gtggccgagt aagagccaga 660
gatctcttgc agtatgtgag cgggcatctt ggcgtcggat gctgggggtca cgttcccagg 720
gatgagaacg tgcagagagc ttcagacctc agcgacaaag agtgttatga cttaagtttc 780

```

&lt;210&gt; 192

&lt;211&gt; 1125

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:297070.1:2000SEP08

&lt;400&gt; 192

```

gacagggtta agcaatTTTT ttaaagcact acatatTTat tgatagaata tctctccagg 60
cagtattcac atgctTTaaa gtgaagctcg tagcttggca gtgcagTTta agcaaataaa 120
ctcaggctgt tgaacaCaat gattaacagg aagcagggga ggggacatag ttactagact 180
ctggccggga caagaggctg aggttaaggt cttagagccg atgacttgca gaatggTcac 240
ctacccact tgcaatTTca tggTgggaac aggtggaccc cctggaatca gaacctctct 300
gaggacatct gTTTTgtgt agacacaggt tgcaggTtag caggagaaca ggcaagccaa 360
atgcaaagga gccactTcag aaatgtgtca cagaaaagtg aaaatgcaac ctagtggtaa 420
gtgaagaggg gaagaagaaa gaaaaaggac cagaaccgtg aactgaaggg acaggggaaca 480
gccagacgag agctTcagcc atcacgagga tgattTcgga acctggagaa aatgtaagtt 540
aaatatatct acactctgat cctatctcaa gagagagata ttttactcat ttcttggtt 600
tgaatgatgg gctctTgga gcactgtctt tttagcgct ctcttatctc tgccctgatt 660
tttgattttg tttacaatac tgagttatgg gagaataaac gttttctgag ggcagctctg 720
tccaatgctt cactgttagc agaagcctgt catcagattt ttgaggggaa agttttttac 780
ccaacagaaa atgcattgaa aactaccctt gatgaagcta cctgctatga gtacatggtt 840
cgaagccact atgtaacaga aaactctctt gaagaagagg ctgggttccc tttagcttac 900
acagtgacca tccacaaaga ctttcggcac ttttgagaga gctctTcagg gcgatttata 960
tgtccccaga atgtctactg tgtgcacctg gatcagaagg cgacggatgc ctttaaaggt 1020
gcagtgaatc agttactcag ctgcttccca aatgcttttc tggcttccaa gaaggagtcg 1080
gttgctctatg gggggatctc caggctccag gctgacctga actcc 1125

```

&lt;210&gt; 193

&lt;211&gt; 640

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1085041.1:2000SEP08

&lt;400&gt; 193

```

cttgctttgc ccacgttgct ggcctcaggc tcagacacct gctctactcc aagcaaatgg 60
ctgctcttcc aatgctgtgg accgggctgg tctcttggg tctcttggga tttccacaga 120
cccagccca gggccatgac acagtgcagc ccaactttca acaagacaag ttcttggggc 180
gctggtacag cgcgggcctc gcctccaatt caagctgggt ccgggagaag aaagagctac 240
tgtttatgtg ccagacagtg gttagctcct ccacagaagg cggcctcaac ctacactcta 300
ccttcctaag gaaaaaccag tgtgagacca aggtgatggT actgcagccg gcaggggttc 360
ccggacagta cacctacaac agccccact ggggcagctt cactccctc tcagtggtag 420
aaaccgacta cgatgagtac gcgttcctgt tcagcaaggg caccaagggc ccaggccagg 480
acttccgcat ggccaccctc tacagcagag cccagcttct gaaggaggaa ctgaaggaga 540
aattcatcac ctttagcaag gaccagggcc tcacagagga ggacattggt ttcttgcccc 600
aaccggataa gtgcattaca agagtaaaca caggtgagag 640

```

&lt;210&gt; 194

&lt;211&gt; 578

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1071544.1:2000SEP08

<400> 194  
gtggcgtctc gctgtcccgc cccgcgcccc tcctctccgc gcgcgggctg gtggctcgttc 60  
tcgtctcttg ttctcttcgc tctccgctgg caccagccga cccaccagct ttcggaccat 120  
ggccaacctc gagcgtacct tcattgccat caagccagat ggcgtgcagc gcggcctggt 180  
ggcgagatc atcaagcgat tcgagcágaá ggggttccgc ctgggtggcca tgaagttcct 240  
tcgggcttct gaggaacacc tgaagcagca ctatatggac ctgaaagacc gtcctttctt 300  
cccggggctg gtgaagtaca tgaactcagg gcctgtgggt gccatggtct gggaggggct 360  
caatgtgggtg aaaacaggcc gagtgatgct gggggagacc aatccagctg attcgaagcc 420  
aggcaccatt cgtggcggat ttctggcatt caagttggca ggaacatcat tcacggcagt 480  
gatccagtgg agagtgccga gaaagagatc gggctatggg ttaagcccga agaactgatt 540  
gactataagt cttgtgcccc tgactgggtg tatgagta 578

<210> 195  
<211> 1816  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:2052480.1:2000SEP08

<400> 195  
ctcgagccgt cagtggcctg tacctggggg agcttgtcag gcttatcttg ctgaagatgg 60  
ccaaggctgg cctcctgttt ggtgggtgaga aatcttctgc tctccacact aagggcaaga 120  
tcgaaacacg gcaactgtggc tgccatggag taagtataaa aagaagagcc ttgctaatac 180  
aagagagatc cgtgggtggac catgggtctg gaaccgctct gaggtgact gcattgccgt 240  
ccagcatgtc tgtaccatcg tctccttcgc ctccggccat ctctgtgcag cagctctggc 300  
ggccatcctg acacgcctcc gggagaacaa cgaagggtgga acggctccgg accacagtgg 360  
gcattggacgg caccctctac aagatattac cctcagtagc caaaacgcct gcacaagggt 420  
ggtgagggaaa ctgggtcccaa gctgtgatgt ccgcttcctc ctgtcagaga gtggcagcac 480  
caaggggggac gccatggtga ccgcgggtggc ctccccgcgt gcaggcccag cggaagcaga 540  
tcgacagggg gctgggctttg ttcccagctg accccgagag cagctccgtg gacgtgcagg 600  
cccaagatgc gggctgaagc tggagtatgg tactgaagaa gaagagccac gggctggcca 660  
cggtcaggat gctgcccacc tacgtctgcg ggctgccgga cggcacagag aaaggaaagt 720  
ttctcgccct ggaactttggg ggaaccaact tccgggtcct cctggtgaag atcagaagtg 780  
gacggagggtc agtgccaagt tacaacaaga tcttcgccat ccccttgag atcatgcagg 840  
gcactgggtg gagctcttt gatcacattg tgcagtgcac cgccagactt ccttggacta 900  
catgggcctc aagggagcct ccctacactt tgggcttcac attctcaatt tcctgcagg 960  
gcagatgagc cattgacaag ggaacactca taggggtggac caaagggttc caagccactg 1020  
actgtgaagg ggaagtgcgt ggtggacatg ctccagggaag ccatcaagag gagaaacgag 1080  
tttgacctgg acattgttgc agtcgtgaat gatacagtgg ggacctgat gacctgtggc 1140  
tatgaagatc ctaattgtga gattggcctg attgcaggaa caggcagcaa catgtgctac 1200  
atggagggaca tgaggaacat cgagatgggt gaggggggtg aagggaagat gtgcatcaat 1260  
acagagtggg ggaggatttg gagacaatgt ctgcatagat gacatccgga ccgatacga 1320  
cacggagggt gtaggagggg tccttgaatc ctggcaagtc agagatacga gaaaatgacc 1380  
agtgggtgat gtacttgagg ggagattgtg cggtcagtat ccatgatcga gccttgatcc 1440  
aagcagggtc tcctcttacc gagggcagat ttcagagcgt ctccggacca ggggcatctt 1500  
cgaaccaag ttctgggtca acagatcgaa cgcgatcggc tggcccttct ccaggtcagg 1560  
aggattctgc agcagctggg cctggacagc acgtgtgagg acagcatcgt ggtgaaggag 1620  
gtgtgcggag ccgtgtccc gcgggcgggc cagctctgcg gtgctggcct ggccgctata 1680  
gtggaaaaaa ggagagaaga ccaggggcta gagcacctga ggatcactgt ggggtgtggac 1740  
ggcagcctgt acaagctgca ccctcacttt tctagaatat tgcaggaaac tgtgaaagaa 1800  
ctagccccctc gatgtg 1816

<210> 196  
<211> 565  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:450105.1:2000SEP08

<400> 196  
gccgtcgcca gctcggccgc cgtactcccg ctaccgagct aggaggcatc accttcgccc 60  
atccaacatg ggtaagacac gtggtatggg agctgggccc aagctcaaga cccacagaag 120  
gaaccagagg tgggctgaca aagcctacaa gaagagccat ctccggcaacg agtggaagaa 180  
accctttgct gggtcattct acgcaaaggg aatcgtcctt gagaagatcg gcattgaggc 240  
taagcagcct aactctgcta tccgtaagtg cgctcgtgtc cagctgggtga agaacgggaa 300  
gaagattgct gcctttgtgc caaacgatgg ttgcttgaac tacatcgagg agaacgatga 360  
ggtgctgatt gctggattcg gtcgtaaggg ccatgctgtg ggagatattc ccggcgctccg 420  
tttcaaggct gtgaaggctt ctggcgcttc cctcctcgct ctcttcaagg agaagaaaga 480  
gaagccgaga tcttaaaacg ctgcaagggt ttgggcctgg tggcgaccc caccctgta 540  
gcgctcacac aagccatagc agacc 565

<210> 197  
<211> 501  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:450581.1:2000SEP08

<400> 197  
ctcgagccgc gccgcccgcg ccccaactcg ccttgcggtga ctgaaagctc gtcgggttcc 60  
gtccacgcga gaagcgagag catggacacc caggtgaagc ttgctgttgt ggtgaagggtg 120  
atgggcagga ccggctccag gggtcagggtg acccagggtca gagttaagtt ctgggatgac 180  
cagaaccggc tcatcatgag gaatgtcaag gggcccgctc gcgagggtga catcctcacc 240  
ctgctcgagt ccgagagggg gcccaggagg ctgcgctgaa gccctagcgt tcttggtggg 300  
ctcttggttg gtctttggtg tcttcgaatt tgtgctatgt ttgaacttct gtgacctgt 360  
cttgagagta gtaagcctgg tgataagatt gcagttttgt ttatcctggt tgaaaatgga 420  
attggtgttg agacatgggt atccttgcta tattatctat aaaaacgacgc atgcattgaa 480  
taccaaaaaa aaaaaaagg g 501

<210> 198  
<211> 569  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:450887.1:2000SEP08

<400> 198  
cgtcggttca cttctccagg aaagggttcg tactcatggc gccgcccag ccaaagtcgg 60  
gcctcttcgt tggcatcaac aagggtcatg tcgtcaccac gccgagctg cctccccgcc 120  
cgtgccaccg caaggggaaa tcaacgaaga ggggtgtctat ggtcaggggc ctgatcagag 180  
aggttgctgg gtttgctcct tatgagaagc gtatcactga gcttctgaag gttggcaagg 240  
acaagcgtgc cctgaagctt gctaagagaa agcttggaac tcacaagagg gcaaagaaga 300  
agagagagga gatggcgggc gtcctcagga agatgaggtc ggctggtaac cactgaca 360  
aaaagaaata gagagcattt caagttcatg gagctggctg ccagagatta tgttcagtg 420  
tctgattttc catacatgta gaaccttaata gacatgtcaa agtattatgt atcgaaccag 480  
ctcatgggat tttgctcctt ccaatgcac caggggttat gtatcgaacc aatttatggg 540  
atcttgctct tattctaag catccatgg 569

<210> 199  
<211> 336  
<212> DNA  
<213> Homo sapiens

<220>

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:460809.1:2000SEP08

&lt;400&gt; 199

```

gcgcacatgggt ggagtgaggct tcaagaagtg cccctcgggc acacagagag atccagaaat 60
tcgccatgaa ggagatgggg actccaaatt tgcacattga tgtgaggctc aacaaagctc 120
tctgggccaa aggaataagg aatgtcccat accatatcca tatgaagttg cccagaaaac 180
ttaatgagga tgaagattca ccagacaagc tctatgcttt ggttcctaca tataacctgt 240
accactttca caaatctata gacaggcaat gtggaagaga gctaaccact gatggttcaa 300
tacattaagt aaaattattt ttaaaaaaga aattta 336

```

&lt;210&gt; 200

&lt;211&gt; 823

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:452089.1:2000SEP08

&lt;400&gt; 200

```

ttacacgtcg tgactgggaa aacaccgtcg tccgcccgcg ccgaaggacg gaaggagaag 60
agggtacggc cgtctcctcg ccccatggc ccacgagaag aagctgtcca acccgatgcg 120
ggagatcaag gtgcagaagc tcgtcctcaa tatctccgtc ggggagagcg gcgaccgtct 180
caccgcgcgc gcaaagggtg tcgagcagct cagcgccag acccccgctc tctccaaggc 240
gaggtacacg gtgcggctcg tcggcatccg gcgtaacgag aagatcgct gctacgtcac 300
ggtgaggggc gagaaggcca tgcagctgct tgagagcggc ctcaaggtea aggagtacga 360
gctgctcagg aggaacttca gcgacaccgg gtgctttggg ttcggcatcc aggagcacat 420
cgaccttggg atcaagtacg atccttcaac aggcattctac ggaatggact tctacgtcgt 480
gctggagcgt gcgggctacc gtgtggcacg ccgcaggagg tgcaagtccc gcgtcgggat 540
tcagcacagg gtgaccaagg aggaactccat gaagtgggta ccaggtaaa gtacgaaggc 600
gtcatcctca aacaaggctc aggccaaacac gttgtaaacc tcacctgtgg gcaaaacagc 660
tctctggtct cttctcctcc tccgtgtcaa cgcaagacca ccaccacct cgccaggctt 720
ttttgggttt aatttgtgtt ctcagacctg atgatttagt ctgctagcac tctgtgggat 780
gctgttgcg caagatcacc accatgtcaa attagttccg tgc 823

```

&lt;210&gt; 201

&lt;211&gt; 779

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1099416.1:2000SEP08

&lt;400&gt; 201

```

tatagggttaa acatccaatt aaaaggcaga tattggctat ccttggggag gccagcaaat 60
ttccggataa gggagactgg tgggcaggaa aaaggcatgc aatggccctc atgagccaga 120
agaatagctg ctgactgatg cctgggggac caagcccata attgtggagc tggagccgca 180
ctgtgtccct tgcttcagtc tgatctttgg ggagcctgca cagttctgac ctgagttaag 240
agctaaggga agagatttgg gaatgtttgg ccctgtgtaa cagcaaaagg attcatttgt 300
gaggggtgtct ggcaggatct cgtaaagata ataagatgaa gggaacatcg ccgtttggaa 360
agtgtcgtga tatgatacac aagttgtgct gcctctgtgg ctctaaggca taccaccttc 420
agaagtcaac ctgtggcaaa tgtggctccc ctgccaagcg caagagaaag tgtaactgga 480
ctgccacggc taaaagaaaa taccacggcg actgggtgaa tgaagcacct aaacattgta 540
tactgcagat ttaggcatgg attctttgca ggaacaacac ctacacccaa gagggcagca 600
gttgtgtcat ccagttcatc ttaagaattt caatgattag tcacacaata aatattccgg 660
tttttaaaaa tgtatatatt ttaaacatat atatgtttat atgtatatgt tatatctgta 720
ttacatatat gtgaaaagag gcagagattg tcagattgga ttaaaaagct gtctgtaag 779

```

&lt;210&gt; 202

&lt;211&gt; 480

&lt;212&gt; DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:255713.1:2000SEP08

<400> 202

```
ctgggattgc aggagtgagc cattatgccc ggccaagtca gggcttttat agctaagatg 60
cctactacct aaaacatgta agcaacttgg cctcctgcaa agaaggatgg tgagaagaat 120
ggccattatg ccatcaacta ggtgggtgacc caagagtaca ccattaatat tcaacaagtgt 180
atccattgag tgggcttcaa gaagtgtgcc cctcaggctc ttaaagagac cggaaatttg 240
caatgaagga gatagaaaac tccagatatg tgcattgata ccaggctcaa caaagctgtc 300
tgggccaag tagtaaggaa tgtccatact gtgtccatgt gcagtggatg gtccagaaaa 360
cataatgagg ttagaaattc accaaataag ctctatactt tggttatata cttgttacca 420
ctttcacaaa tctacagtca gtgtggatgg gaactaactg ctgatcatca aatacatcaa 480
```

<210> 203

<211> 604

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:998903.1:2000SEP08

<400> 203

```
gcccggtgct ccctctctgc agatctccct tcgcagccgc cgaggaaccc cttcacagcg 60
cggaagatgg ctgaacagga ggccccagtc gcggttgagg caccaacccc agttcttggg 120
gagcctatgg acctgatgac tgctctgcag ctcgatgaaga agaagtcaag tgctcatgac 180
ggtcttgtga agggctctccg tgaggctgcc aaggccatcg agaagcacgc cgctcagctt 240
tgctgtgctt ctgaggactg tgaccagcca gattacgtca agttggtgaa ggcactctgc 300
gctgagcaca atgttcacct ggctactggt cctgccgcta agactcttgg cgagtgggccc 360
gggctttgca agattgactc tgagggcaag gcaaggaagg ttgtaggctg ctccctgcgc 420
gtcgtcaagg actacgggtga agaactctgag ggccttaaca tagtgcagga gtatgtcaag 480
tcgcactaga tgtgtgacat gtttcagtga tactcttgat ttggacctgg gcttaaaatt 540
atgcttttgc atctggggct tgccctaggag attagaacat ttactagaag agcgaagaga 600
tatg 604
```

<210> 204

<211> 617

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1119656.1:2000SEP08

<400> 204

```
aaaaataaaa aaataaaaaat aaaattttta tgtaagaaaa ttgaaggaaa aaatggactt 60
tgtgttatgt cttcccttac ttgttctgtc tgtatatattcc tatattgatt cggtagtagc 120
aatgcataat tttccacaaa acttcatcct tacatgtaga gaaagtgttg ataaaaatatt 180
ctgcaataaa gtgctctttg caaatatgta cttcattttc actgtttatt caattttctt 240
aataccatat aagttcttac aggagagctt cagattctcc atacaaaatg gttaatctaa 300
aactatgctg gatagagtaa aacccttgct gagcttctta gaaaaaaaaa aaagatgctt 360
tatgtgtcaa atataggcta tgttaaatca tagcaaagcc tttggcctca tggcactccc 420
gaaagcaaga tggggtcacc agcagcccta ctgggagcca cccacggaag ctggggcagg 480
ttctcactct ggctcacatct gcccaaacca gcatagtctg atctggaaac acggcctcaa 540
tatgtgccgc cgggtgttccc atcagtatgc aagggacata ggtttcattg agtgggacta 600
agtgatcttc cttgaat 617
```

<210> 205

<211> 455

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1096907.1:2000SEP08

<400> 205

```
tggagcttcg gttccggcga catggctaaa cgcaccaaga aggtcgggat cgtcggaaaa 60
tacgggaccc gctatggtgc ctccctccgg aaaatggtga agaaaattga aatcagccag 120
cacgctaagt acacttgctc cttctgtggc aagaccaaga tgaagagacg agccgttggc 180
atctggcatt gtggttcctg catgaaaaca gtggccggtg gggcctggac ctacaatacc 240
acttctgcag tcacagtga gtctgccatc agaagactga aggcaggtga gtttgtctgg 300
tggtggagag aaagaggtga ggtgtagctg ggcagggggg ctttctctgc ccctcaccga 360
atccccattc caggctcaag gggcagaagt ccccgatcg cctcagcagt ccccccacagt 420
aggcaataaa agcctgaaaa ggaacgagcc gggca 455
```

<210> 206

<211> 577

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1323741.1:2000SEP08

<400> 206

```
gaagacatat attttcattt ttttctgggc atgtacttgc cgtatcatta ggtaagttta 60
tatttgaatc tacgctcttt ccctcggagc gggcggcgcg gcggcgctcg cggcctgtgc 120
agcaatggcc aagattaagg ctccgggacct gcgcggcaag aagaaggagg agctgttgaa 180
acaactggat gatctgaagg tggactgtc ccagcttcgc gtggccaaag tgacaggcgg 240
cgccgcgtcc aagctctcca agatacgagt cgtacgcaaa tccatcgccc gtgtcctcac 300
tgtcattaat cagactcaaa aggaaaacct caggaaattc tacaaggga agaagtacaa 360
gccctggac ctgcgaccca agaagacaag agccatgcgc cgccgggtca ccaagcatga 420
agagaagctg aagaccaaga agcagcagcg gaaggagcgg ctgtaccac tgcgcaagta 480
cgcagtcaag gcctgagacg acgacaataa caataaagtc caaaactgac gaaaaaaaaa 540
aaaaaagcgg ccctcgggat ctagaaccta aatcgcg 577
```

<210> 207

<211> 443

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1098372.1:2000SEP08

<400> 207

```
tgttgccctc tttctgcctc cgctaccgcc atggcgccca tgaaaaagct tgtggtaaag 60
gaggggctaaa aaaaaaggaa gcagggttcca aagtccaactc ttgatcgcac ccaccccgta 120
gaagatggaa tcatggatgc tgccaacttt gagcagtttt tccaagaaag gatcaaaatg 180
aacggaaaag ctgggaactt tgggtggagg gtagtgacca tcgaaggag caagagcaag 240
accagcgtga catccaagct gcccttttcc aacagggtatt tgaaatatct caccacaaaa 300
atatctgaag aagaataatc tacatgattg gttgcgcgta gttgctaaca gcaaacagag 360
ttacgaatta cgttacttcc aaattaacca ggacgaagag gaggaaaacg aggattaaat 420
ttcatttatc ggccggggcg ggt 443
```

<210> 208

<211> 552

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1006783.1:2000SEP08

<400> 208

```
cagtaactcg ccaaaatgac gaacacaaaa ggaaagagga ggggtactcg gtatatgttc 60
tctagacctt ttaggaaaca tggagtcgtt cctttggcca catacatgcg aatctacaag 120
aagggtgata ttgtagacat caagggaatg ggcactgttc aaaaaggaat gccccataag 180
tgttaccatg gcaaaaccgg aagagtctac aatgtcacc agcatgccgt gggcatcatt 240
gtgaacaagc aagttaaagg caagattctg gccaaagagga tcaatgtgcg gattgagcac 300
atcaagcact caaagagcag agacagcttc ctgaagcggg tgaaggagaa cgatcagaag 360
aaaaaggaag ccaaagagaa gggcacctgg gtccagctga agcgccagcc tgcgccaccc 420
agagaagccc actttgtgag gactaacgga aaggagcctg agctgctgga gccattcca 480
tacgaattca tggcctaata tatacaaagg aaataaagaa cctgaactgc aaaaaaaaaa 540
aacaagcgg cc 552
```

<210> 209

<211> 510

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1097562.1:2000SEP08

<400> 209

```
gaggactagg cggtcggggg agctccgcta gttggtgttt gacgctctgt atcataatcc 60
tcacttctgc cctctgtgta ttctagggtt gggcctgtcc cgcacctaag gcaagaggat 120
ggtggctgca aagaagacga aaaagtctct ggagtcaatc aactctcggc tccaacttgt 180
tatgaaaagt ggaaagtacg tgctggggta caaacagact ctgaagatga tcagacaagg 240
caaagcgaaa ttggttatcc tcgccaaaca ctgtccagct ttgaggaaat ctgaaataga 300
atactatgcc atgttggcta aaactgggtt ccactactac agtggcaata acattgaatt 360
gggcacagcg tgtggaaaat actacagagt atgcacactg gctatcattg acccagggtg 420
ttccgatatt attagaagca tgccagaaca gactggtgag aagtaaaca gaaagttctc 480
ctttaataaa actttgccag agctcctttt 510
```

<210> 210

<211> 868

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:998868.1:2000SEP08

<400> 210

```
catggctgtc gggaagaaca agaggatctc caagggaag aagggtggca agaaaaagac 60
cgctgacctt ttctctaaga aggattggta tgacatcaag gcaccgtcgg tgttcagtgt 120
gcgcaacatc gggaagactc tcgtatccag gacacagggt accaggattg cttctgaggg 180
tctgaagcac agagtctttg aggtttgcct agctgatctt cagggcgacg aggatcaagc 240
ttacaggaaa atcaggctcc gtgctgaaga tgtgcagggc aggaatgtgc tcacaaactt 300
ttgggggatg aatttttacca ctgataaatt gaggtcccta gtgaagaagt ggcagacatt 360
aatcgaagcc catgctgatg tcaagacaac cgacaactac atgttgctgt tgttctgcat 420
cggttcacc aagaggcgcc caaaccaggc caagagaacc tgctatgctc aggcattcca 480
gatccgtcag atccgcgcta agatggttga gattatgatc aaccaggctt ctacatgtga 540
cctgaaagaa cttgtttcga agttcattcc agaagttatt ggtaaggaga ttgagaagtc 600
tacctccagc atcttccac ttcaaaacgt cttcatccgc aagggtgaaga tactcaaggc 660
ccccagttc gatcttggga agctcatgga ggttcacggc gactacaaag aggatgttgg 720
tgtgaagcct gagaggcctg tcgagggaga tgaggctggg caggaggttg ctgctgccga 780
gtagtttgtt tgctagtctt atatctaggc gagtctggac atgaggttgc ttatacgtga 840
actatagagt tgtcttggct atgtgttt 868
```

<210> 211

<211> 1010

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1063383.1:2000SEP08

<400> 211

```
tgcaggtcag ctcactcagt tgtaacaga gattttgac gtcaggagct ctgggaatat 60
aaagatgaaa cttccccctt gccttcaagg atatcatggg ccagaaggca aagttgtttt 120
gaatacgtgg ttcataggag acccactctg tgccaactga tggctgcaaa gagaacagaa 180
ggggtgctgc ttaggaaat aaatgaatgg ctcaagaagac cactactgagg aagttgtgag 240
ttgataatgg aagatctcca ggtttgaggc atccttagag ggatatgatg gttttgtgtg 300
tgttgggggt gtggttagcg agctgctacc tagggaatta gaaggttttc tttattgaac 360
atttaccctg tgacaggtac tgcaggcatt cagcacacaa tgcgtctcc attttacagg 420
tgaggaaact gagactccag ttcaagtaga tggtaaggc cagtactacc ggaaggacca 480
tctgggggtt cagacactgg cagggtggga tttgctgcc cttgcaaatt gagagtgtct 540
tggggtcagt tttgatttgc tcagctgttg gcattctttg ggctctgagt gggtaggtgtg 600
acccttgacc tcctgggatc gcactctggag actgcctagt attctgccag cttcggaag 660
ggagggaag caagcctggc agaggcacc attccattcc cagcttctc cgtagctgtg 720
gattggaaga cactctgcga cagtgttcag tccctgggcg ggaaagcctc cttccaggat 780
tcttcctcac ctggggctac ttcttcccca aaaggcatca tggctgccct cagatccctt 840
gtgaagccca agatcgtcaa aaagagaacc aagaaattca tccggcacca gtcagaccga 900
tatgtcaaaa tcaagcgtaa ctggtggaaa ccagaggga ttgacaacag ggttcataga 960
aggttcaagg gccaggctct gatatgatgc ccagcattgg ttataggagc 1010
```

<210> 212

<211> 408

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:1400567.1:2000SEP08

<400> 212

```
ggtgacccga agaatacacc catcaacatt cacaagtgc tccatggagt gggcttcaag 60
aagcatgccc cttgagcact caaagagatt cggaaatttg tcaggaagga gatgggaact 120
ctagatgtgc acaactgata acaggctcaa caaggctgtc tgggccaag gaataaggaa 180
tgtcccatgc cgaatccatg tgcagttgtc cagagaatat aatgaggatg aagattcacc 240
aaataagctc tgtacttcgg ttacctatgt acctgttacc actttcaaaa atctatagac 300
agtcatttgt gatgagaact aatcactgat tgtcaactat atcaacaaa gttataaaac 360
tgcaaaaaaa caaaaaaaa tgaacccaac ccaggggaaa ccaaagat 408
```

<210> 213

<211> 643

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:449404.1:2000SEP08

<400> 213

```
atccgtcgcc ccacggaagt ctggcgggcg cggcggtgct tgtgatctca acccaagtgc 60
tgccctcgct ccggtcacc gtgctccac gcaaccatgt cgaggaggaa gaccaggag 120
cccaaggagg agaagctcac ccttggaacc actgtccgtg aaggagagta tgtctttgtt 180
gttgctcaca tctttgcac cttcaatgac accttcattc atatcactga tttgtctggg 240
agggaaactc tggttcggat caccggtggc atgaaggtga aggctgaccg tgacgagtcg 300
tcaccttacg ctgctatgct tgctgctcaa gatgtcgac agcgctgcaa ggagcttggc 360
attactgcac tgcacattaa gcttcgtgcc accggaggca acaagaccaa gacccttggg 420
cctggtgccc agtctgccct caggcgctt gctcgttccg ggatgaaaat cggacgcatt 480
gaggacgtta ccccggtccc caggacagc actgcagaa agggcggtag gaggggaagg 540
aggctgtagg cgtctcttct gcgtgccatt ttgctggctc ttgtcgggtg gccatcatga 600
```

agctagtcctc tccccctggg ttggtgttgc tgttttctgt gtt

643

&lt;210&gt; 214

&lt;211&gt; 780

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:449941.2:2000SEP08

&lt;400&gt; 214

gcggcgccgc gggcgagcaa ggagcaggcg gaggaggagt tgtgccaagc cacaatgggt 60  
atctcgctgt actcgatgca caagcgccga gccaccggtg gaaagcagaa ggcctggagg 120  
aagaagcgaa agtatgagct tggtcgccag ccagccaaca ccaagttgtc aagcaataag 180  
acagtgagga gggctcgtgt tctgtggagg aatgtgaaat ggagggtctc tgcgcttgat 240  
actggttaact actcatgggg aagtgaagct gttacccgca agaccggtat cctcgacgtg 300  
gtctacaatg catcaaacaa tgagcttgtg agggacacaaa cccttgtgaa gaggtgccatt 360  
gtgcaagttg atgctgcccc attcaagcag tggtagctca ctactatgg agttgacatt 420  
ggtaggaaga agaaaacccc tgctgcaaag aaggataatg ctgagggaca agaggttgag 480  
gcagcagctg aggaaacgaa aaagagcaac catgtcacga ggaagcttga gaagcgcaag 540  
gagggacgta cccttgaccc acacattgag gagcaatatt ggcagtgagc ggttgctggc 600  
atgcatttct tcccgccttg gacagtgtgg ccgagctgat ggttacatcc ttgagggtaa 660  
agagcttgag ttctacatga agaagctaca gaggaagaag ggcaagagcg ctgtagctta 720  
gagcattctc atggattgtc ttgcaccgaa atgtgtgtcc tgatgagtag actttttttg 780

&lt;210&gt; 215

&lt;211&gt; 715

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450229.1:2000SEP08

&lt;400&gt; 215

ctagcgccgg cgctccact tctgttttct cactctctcc tccagctcag ggtccggcgg 60  
cgaaggggaag gcaagatgta caccgcgagg aagaagatcc agaaggagaa ggtcttgag 120  
ccctccgagt tcgaggactc cgttgcccag gctttctttg atctggagaa cgggaaccag 180  
gagctcaaga gcgacctcaa ggacctgtac atcaacaatg ctatccagat ggatgttacc 240  
gggagttaga aggtctgtgt cattcacgtc ccataccgcc tgcgcaaggc cttcaggaag 300  
atccatgtca gactcgtcag ggagctggag aagaaattca gcggcaagga tgtggtaatt 360  
gttgctacac ggaggattgt gaggccaccc aagaagggtt cagctgttct gcgcccctgc 420  
accaggactc tgaactgctg tcacgatggc atcttgaggg atgttgctta ccagctgag 480  
attgtgggga agcgtgtcag ataccgtctg gatggttcca agattatcaa gattttcttg 540  
gacccaaagg agaggaacaa cactgaatac aagctggaga cctgcactgc ggtctaccgc 600  
aggctgtgtg ggaaagatgt ggtctttgag taccctatga ccgaaaatgc ataaatatga 660  
tgccctctgg atatctccac tctatttctg tgattctgaa tgttatgttg ggtct 715

&lt;210&gt; 216

&lt;211&gt; 693

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450399.3:2000SEP08

&lt;400&gt; 216

ctccgcgcgt cctctccgct tccaagatgt cgaagcgagg acgtggagg tctgctggta 60  
acaagttccg gatgtcactt ggtctaccag tggcagccac tgtgaactgt gctgataaca 120  
ctggagccaa gaacctttac atcatttctg tgaagggaat caaggggcgc cttaacaggc 180  
ttccttctgc ctgtgttggt gacatggtta tggcaactgt gaagaaggga aagcctgacc 240



```
ttaggaagaa ggtgatgcca gctgtcattg tgaggcagcg caagccatgg cgccgaaagg 300
atggtgtcta catgtacttt gaagacaatg ctggagtgat tgtgaaccca aagggagaga 360
tgaagggtc tgccatcact ggacccatcg gaaaggagtg tgctgatctg tggcctagga 420
ttgccagcgc ggcaaatgcg attgtctagt ggaatcagtt gatcgaaatt tacttcaaat 480
aacatttgtt tgtgttgcca attgggaaga gagtttttag ttagacgagc ttttgatggc 540
cctaaagatg ttttcccgga tgatgtttgt gctgctggac cgtttcctgt tagcttcgat 600
ggcatcctcc ccatatcgat aaactgttgc ttagcctgac aatcggaaga gtttaaattg 660
taaagatttg atatctccaa aaaaaaaaaa agg 693
```

&lt;210&gt; 217

&lt;211&gt; 568

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:455771.1:2000SEP08

&lt;400&gt; 217

```
ccgcttcttg cagccgtcgc cgcaagcgc tccagagaac ttccgtcaac atggggaaga 60
cacgtgggat gggagccggg cgcaagctca agaccaccg caggaaccag cgggtgggctg 120
acaaggcata caagaagagc catttgggca atgagtggaa gaaacccttc gctgggcat 180
cccatgccaa gggcattgtc ctggagaaga ttggtattga ggccaagcag cccaactccg 240
ctatccgtaa gtgtgctcgt gttcagcttg ttaagaatgg caagaagatt gctgccttcg 300
tgccaaatga cggttgtttg aactacattg aggaaaatga tgaggctctg attgctggat 360
ttggctgtaa ggggcacgct gtgggagata ttcttggtgt ccggttcaag gtcgtcaagg 420
tttccggtgt gtctctgctt gcccttttca aggagaagaa agagaagcca aggtcttaga 480
ttgctcttgc taccaaaatc agcaagcgtg gagttgaaac gggagggcgt tagatgatta 540
agaagaatgg ttgcttgcta tgtttgca 568
```

&lt;210&gt; 218

&lt;211&gt; 564

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:720459.1:2000SEP08

&lt;400&gt; 218

```
gtccttgcta cgtcacaaa aacagttggt ggggacaaga acggtggcac ccgggtggtg 60
aagcttcgaa aaatgcctag gtattaccct actgaagacg tgcctcgga gctgctgagc 120
cacggcaaga agcccttcag ccagcacgtg aggaggctgc gctccagcat cactccggg 180
actgtcctga tcacctcac tgggcgccac aggggcaaga gagtgggttt cctcaagcag 240
ctgggcagtg gcttgetact tgtgactgga cctcttgccc tcaacagagt tcctctgctg 300
aggacacacc agaagtttgt catcgctacc tctacaaaag ttgatatacag caagggttaa 360
attcccaaac acctgactga tgcttacttc aagaagaagc cacttcgcaa gccagggcat 420
caggaggggtg agatcttcga cacagagaag gagaaatacg aaattacaga gcagcgaaag 480
gctgatcaga aagctgtgga ctgcagatt ttgccaaaga tcaaagctgt cccccagctc 540
cagggtctacc tgcggtctca gttc 564
```

&lt;210&gt; 219

&lt;211&gt; 533

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:723156.1:2000SEP08

&lt;400&gt; 219

```
tacgccgcaa cctccctcga gcccgagcaa acaagggcgg acgcggagcc ggccggcgag 60
cggagatggc ggagaagaag cgaggagccg gcactcgcaa ggatgaagtg gtgacacgcg 120
```

```

agtacacccat caacctccac aaacgcctcc acggatgcac cttcaagaag aaggcaccac 180
atgccatcaa ggagatcagg aagttcgctc agaaggccat ggttactact gatgtccgga 240
ttgacgtgaa gctcaacaag cacatctgga gcagtggtat ccgcagtggt ccgaggcgtg 300
tccgtgtcag aattgcccgc aaccggaacg atgaggaaga cgccaaggag gaactctact 360
ccctgggtcac agtagctgag atcccaccgg aggggtctcaa aggtttggga actaagggtg 420
ttgaggatga ggattaagcc ctgtatacca tgagataaat ttgagaatgt tgcttcagct 480
tattgttata tactgccaat gactcatgta atagaagctt cgagacttgt ctc 533

```

&lt;210&gt; 220

&lt;211&gt; 504

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:728055.1:2000SEP08

&lt;400&gt; 220

```

ggcgcggcgc ggggtgaggag ctctgtttca tacttgtagc tatgccgtcc aagggtccgc 60
tgcagtcctg tcaagtcttc ggacgcaaga aaacagccac agctgtggcc cactgcaaac 120
gaggaaatgg gctcatcaag gtgaatggac gtcccctgga gatgatcgag ccgcgcacgc 180
tgcagtacaa gttactggag cctgtttctgc ttctgggcaa ggagagattt gctgggtgtg 240
atatccgggt ccgtgtgaag ggtggtggtc atgtggccca aatttatgct atccggcagt 300
ctatctcaaa agctctgggt gcttattacc aaaaatatgt ggatgaagcc tctaagaagg 360
agatcaaaga tatctcatc cagtacgac ggaccctgct ttagctgac ccctgtcgct 420
gcgaatccaa gaagtttggc ggtcccgggt cccgtgccgc gtaccagaaa tcctaccgat 480
aagccagtct caaggatcag gggt 504

```

&lt;210&gt; 221

&lt;211&gt; 442

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1020789.1:2000SEP08

&lt;400&gt; 221

```

ggcctacacg ccgcccgttg tgccgcgcgc atgtccctag tgatccccga gaagtttcag 60
cacatcctgc gactactcaa caccaacatc gatggggcgc ggaaaatagc cttcgctatc 120
actgccatta aggtgtgtgg gcggagatac gctcatgttg tttgaggaa agcagacatt 180
gacctcacca agagggctgg ggagctcacg gaggacgagg tggagcgtgt gatcaccatc 240
atgcagaacc cagcacaata caagatccct gactggttct tgaacagaca gaaggacgtg 300
aaggatggga agtatagcca ggttctggcc aacggtctag acaacaagct gcgtgaggac 360
ctggagcggc tgaagaaaat ccgagcccat agagggtgc gccacttttg ggccttcgt 420
gtccggggtc agcacaccaa ga 442

```

&lt;210&gt; 222

&lt;211&gt; 498

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1071728.1:2000SEP08

&lt;400&gt; 222

```

ctgtggagcc ttgtcaagtc tgctctgtcc tgactcaccg ctgttcgctc ctgetgagga 60
acaagtcggt caggaagcta cgctcagcc atggcattta aagataccgg gaagactccc 120
gtggagcccc aagtggcgat tcaccggatt cgaatcacgc tcaccagccg caacgtgaag 180
tcgctggaaa aggtttgtgc ggacttgatc agaggcgcaa aggagaagaa tctgaaagt 240
aaaggaccag tgcgcatgcc taccaagaca ctgagaatca ctaccagaaa aacccttgc 300
ggtgaagggt ccaagacgtg ggatcgtttc cagatgagaa tccacaagcg actcattgac 360

```

ttacacagtc cttcagagat tgtaagcag attacttcca tcagtattga gcctggagtg 420  
 gaggttgaag tcaccattgc agatgcctaa gtcaactgtt taaataaatt gacttaattg 480  
 ttaaaaaaaaa acaaaaca 498

<210> 223  
 <211> 522  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1084329.1:2000SEP08

<400> 223  
 ggaagctcct gccctccca aagccgaagc caaagcgaag gccttgaaag ctaagaaggc 60  
 agtgctgaaa ggtgtccaca gtcacaaaaa gaagaagatc cgaacgtcac ccactttccg 120  
 gcggcccaag accctgcggc tccggaggca gccaaaatat cctcgaaaga gtgcacccag 180  
 gagaaacaag cttgaccact atgctatcat caaattccca ctgaccaccg agtcagctat 240  
 gaagaaaata gaggacaaca acacgcttgt gttcattgtg gatgttaag ccaacaagca 300  
 ccagatcaaa caggccgtga aaaaactcta tgatatagat gtggccaaag tcaatactct 360  
 gatacggcct gacggagaga agaaggcata tgttcgcttg gctcctgatt atgatgctct 420  
 agatgttgcc aacaagattg ggatcatcta aagtgaagtc agatggttaa ttctaaatat 480  
 atactttttt tccacaaaaa aaaaaaaaaa agcggccgcc ga 522

<210> 224  
 <211> 442  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:246422.1:2000SEP08

<400> 224  
 tggtgccttc tttctgcctc cgctaccgcc atggcgccca tgaaaaagct tgtggtaag 60  
 gagggctaaa aaaaaaggaa gcaggttcca aagttcactc ttgatcgcac ccaccccgta 120  
 gaagatggaa tcatggatgc tgccaacttt gagcagtttt tccaagaaag gatcaaatg 180  
 aacggaaaag ctgggaactt tggtggaggg gtagtgacca tcgaaggag caagagcaag 240  
 accagcgtga catccaagct gcccttttcc aacagggtatt tgaaatatct caccaaaaaa 300  
 tatctgaaga agaataatct acatgatttg ttgcgcgtag ttgctaacag caaacagagt 360  
 tacgaattac gttacttcca aattaaccag gacgaagagg aggaaaacga ggattaaatt 420  
 tcatttatcg gccgggcgcg gt 442

<210> 225  
 <211> 601  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1086066.1:2000SEP08

<400> 225  
 gccagaactg cgcgtggtcc gcgccgatcg actgagaagc ccggtttgcg ctctcagaat 60  
 gactgaatgg gaaacagcca caccgcgggt ggcagagacc ccggacatca agctcttttg 120  
 gaaatggagc actgatgatg tgcagatcaa cgatatctct ctacaggatt acattgctgt 180  
 gaaggagaag tatgccaagt acctgcccc aagtgacgga cggatgctg ccaagcgttt 240  
 ccgcaaagca cagtgtccca tcgtggagcg ccttactaac tccatgatga tgcacggtcg 300  
 taacaacggc aagaagctca tgactgtacg aattgtcaag catgccttg agatcatcca 360  
 cctgctcact ggtgagaacc ctctgcaggt cctggtgaat gctatcatca acagtggccc 420  
 ccggaagac tcaacacgca ttgggcgggc tggaacagtg agacggcagg ctgtggatgt 480  
 atccccactt cgccgagtga atcaggccat ctggctgctg tgcatcgagg gctcgtgagg 540  
 ctgctttccg gaacatcaag accatcgctg agtgccttgc acgatgagct cattaatgct 600

g

601

&lt;210&gt; 226

&lt;211&gt; 1207

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:223142.1:2000SEP08

&lt;400&gt; 226

```
ggcctgagtt gatctatgaa cattattcac tatgaccaca caagaggaca ggttttcttt 60
ggtcttttaa aggaactgat taaagaaaac ctgtcctctt gtgtggcctg agttgatcta 120
tgaacattat tcaactatgac ccagaaaatc ccacgaagtc gtgcatgtta gaaggatgaa 180
accctcatgt tcaactgtaaa ctgtgaagggt gtcaggcca tcacagggtat gcctatatga 240
aaagccacca agtatctgaa atattttctt ttacagaaat tatgtgtacc atttcagtct 300
tagaaagtgg agttggtcgc tgtacgcaag acagacattg gggctggact caccatcagt 360
ggccccgaaa aggtactgaa atttgcttac aagtgcagag ttatgctgaa cttaagggta 420
tagatgtaga ttccctgggt attgagcaca tccagggaaa aggcacccat aatgtaccac 480
ctgacttaca gaactcatgg gcagatgaac ccattccaca actcccctgc cacatccaga 540
tgaatgcttag tgaaaagaaa caccttggtc caaaagcaga aaaggaggat gcacgggaaga 600
aaaagatacc ccagaagaaa cataaactta agagacaaac aaattcagcc aaaagaaaat 660
gcaataaaaa gtcaaaaaca aaacaaacaa aaaaacctct aaaattcata aaacacaaaa 720
tgacaaactc tattaatatat gccctatga tatctcttag ataaaagtgt gccgtaatag 780
agaaaaatgaa gataggtggt ggtgactggc agattctggg tctacgatga atcagagaag 840
gatgggtttc actttgtttt acattttttt aatgcttgat gttgacttgg tagaaagatg 900
ataataaaga gattcttctg tctatgtgtt tcttttagat tactcatgtg gactgaagaa 960
aaaaaattaa atttgtaaaa gaaagtgagt ttgatttagg attgtgactg aggacaatag 1020
cctggaagca gttctgacag atggctgcta tgtaatatat tgggttcacag ttataagag 1080
acatgtgggc cctgtgctct atcttgtttt tttcttcaa gcataattct ggagagctgg 1140
agagtcccag agtcaggaat gttataaagt caggctggaa accagaaata agtaacatga 1200
ctaaaca 1207
```

&lt;210&gt; 227

&lt;211&gt; 680

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:885368.1:2000SEP08

&lt;400&gt; 227

```
gcctctcttc ctccagagct cgctgccacc ggccaccgcc accctcgatc acgccgtccg 60
tcgaggtaag gtcagcatgg tgagggtcag tgtgctcaac gatgcgctca agtccatgta 120
caacgctgag aagatcggca agaggcaggt catgatcagg ccgtcgtcca aggtcatcat 180
caagtctctg acggtcatgc agcgccacgg ctacattgga gaggtcgagt acgttgatga 240
ccaccgatcg ggcaagatcg tggtcgaact caacgggagg ctgaacaagt gcggcgatcat 300
cagccctcgc tttgatatcg gcgtgaaaga cattgaggga tggactgcaa ggctgctccc 360
gtccaggcag ttcggataca tcgtcctcac aacttcggca ggcacatgag accacgagga 420
ggcccgccgg aagagcgtag gaggaaggt tctaggtttc ttctattgaa gaagctgaga 480
atctttgtta tggcctgcct gcctgtctta ctctctcaga ctagagtacc tgttatTTTA 540
ctttaatatg tctctctcag actagagcgc cttcttttac ttgaatattg gtgactgcat 600
tggtctttca ctttaaaatc ggtgactgca ttgatcgaat tatctgagcc ggtcttcaga 660
tacctagtgg gcacaacctt 680
```

&lt;210&gt; 228

&lt;211&gt; 566

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:481782.1:2000SEP08

&lt;400&gt; 228

```
ggcggagcgt aaggcgaagg agtagcagca gcaggcggcg ccgagtagcg gctcccatc 60
tcgagcttgc caccatggct agaggattga agaagcattt gaagaggctc aatgccccca 120
agcattggat gctggacaag cttggcggac gcttttgctc ccaagccatc ttctggacct 180
cacaagtcca gggagtgcct tcctctgata ctcacatca ggaacaggct caagtatgct 240
ctgaactacc gtgaggctcat ttctatcctg atgcaacgcc atgtacttgt tgatggcaag 300
gtcaggacag acaagaccta ccctgctggg gttcatggat gtcatttcca tccccaaagac 360
caatgagaac tacaggctgc tgtacgatac caagggccgc ttccgccttc acccaatcag 420
ggatgaggat gctaagttca agctttgcaa ggttaggtct gtccagttgg ggcagaaggg 480
catcccttat ctgaacacgt atgacggccg caccatccgt taccctgacc ccctcatcaa 540
ggccaacgac accatcaaga tcgatac 566
```

&lt;210&gt; 229

&lt;211&gt; 480

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1093813.1:2000SEP08

&lt;400&gt; 229

```
gcggcctcat ccgcgaacag cgcagcaatg gccaaagatta aggctcggga cctgcgcggc 60
aagaagaagg aggagctgtt gaaacaactg gatgatctga aggtggaact gtcccagctt 120
cgcggtggcca aagtgcacagg cggcgccgcg tccaagctct ccaagatacg agtcgtacgc 180
aaatccatcg cccgtgtcct cactgtcatt aatcagactc aaaaggaaaa cctcaggaaa 240
ttctacaagg gaaagaagta caagcccctg gacctgcgac ccaagaagac aagagccatg 300
cgccgcggcg tcaccaagca tgaagagaag ctgaagacca agaagcagca gcggaaggag 360
cggtgttacc cactgcgcaa gtacgcagtc aaggcctgag acgacgacaa taacaataaa 420
gtccaaaact gacgaaaaaa aaaaaaaaaa cggccctcgg gatctagaac ctaaatcgcg 480
```

&lt;210&gt; 230

&lt;211&gt; 543

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:449413.2:2000SEP08

&lt;400&gt; 230

```
tccaactcta acccgcgcgc ttccggctcc cgcagcgtcg ccgaccttca cagccgggtc 60
cttcctctgc caccatgggg aagacacgtg gtatgggagc tgggcgcaag ctcaagacct 120
accgcaggaa ccagagggtg gctgacaaag catacaagaa gagccacttg ggcaatgagt 180
ggaaaaaacc ttttgctgga tcatctcacg ccaagggcat cgttctggag aagattggta 240
ttgaggccaa gcagccaaat tcggccatcc gtaagtgtgc ccgtgttcag ctgggtcaaga 300
atggaaagaa gattgctgcc tttgtgccga atgatggttg cctaaactac atcgaggaga 360
atgatgaggt gttgattgct ggatttggtc gtaagggtca tgctgtggga gacattcctg 420
gtgtcagggt caagggtgtt aagggtgtct gtgtgtcgct gcttgcactc ttcaaggaga 480
agaaggagaa gccaaaggtc tagatcactt ttcggtagtc aagaatggtg taaactgccc 540
aag 543
```

&lt;210&gt; 231

&lt;211&gt; 537

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450105.1:2000SEP08

&lt;400&gt; 231

```

gccgtcggca gctcggccgc cgtactcccc ctaccgagct aggaggcatc accttcgccg 60
atccaacatg ggtaagacac gtggtatggg agctggggcg aagctcaaga cccacagaag 120
gaaccagagg tgggctgaca aagcctacaa gaagagccat ctgggctaac gagggaaga 180
aaccctttgc tgggtcatct caccgaaagg gaatcgctct tgagaagatc ggcattgagg 240
ctaagcagcc taactctgct atccgtaagt gcgctcgtgc tccagctggt gaagaacggg 300
aagaagattg ctgcctttgt gccaaacgat ggttgcttga actacatcga ggagaacgat 360
gaggtgctga ttgctggatt cggtcgtaag ggccatgctg tgggagatat tcccgccgtc 420
cgtttcaagg tcgtgaagggt ctctggcggt tcctcctcgt ctctcttcaa ggagaagaaa 480
gagaagccga gatcgtaata cgtgcaagg gtttgggcct ggtggcgcac cccacac 537

```

&lt;210&gt; 232

&lt;211&gt; 764

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:814285.1:2000SEP08

&lt;400&gt; 232

```

ctcagttgtg agggggaagg tagactcctt tccagacctt tgataaagg gctgctcaaa 60
aaagccacga ccttgggcct ctgcttaaaa ccatggcttt aaagtaagga gggaagcagg 120
atatataaag aagaccgtta tccgcgtctt ttctaactgt cttatgtcct gtgagtgtca 180
cccaattaga agtaaaccgc aagtggaagt cagcaggaaa agtaagcatg ggatgatgca 240
gttctttaca ggggtgttga agagaagaag taaggttcac tgctgtgcca gaaaccctta 300
agaacaaagt gagggaattt tcacagaggc tgaagatcaa gcgacctcga cgaataaagt 360
ttagcgccaa aagatgcttc taaaggccaa ggaggaagct tatcttatga aaaagcaaa 420
cacgtatcac aagggaatat atgcagatgt acagaactga aattcaagat atcgaggata 480
gcaagaaaag ctggcaactt ctatgtatct gcagaaccca aattggcggt tgtcatcagg 540
atcggagggt atcactccgg gtggagcccc aaagggtctga aaagggtgtg caaacttctt 600
tgccttcac aattcttcca tgaacaactt tgtgaaagct cagcagggtc tcaataatgt 660
gctgagggtt tgtagaacca tatattgcat gggcatatcc aaatctgaag tcagtaaata 720
aactaatcta caaatgtggt tatagcaaaa aacaaaaaaa aggg 764

```

&lt;210&gt; 233

&lt;211&gt; 522

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1142855.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 113

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 233

```

aatattagta aaacaaccaa aagcccatgt catcatcaga ctctcagat tcttctttgt 60
ttgcttccac tttctttctca gctggagcag cagcagtggg gtgggcagga ccncctgctg 120
gtgcagcacc agctgctgga gcagggtccac cagcccctac attgtagatg aggtctccaa 180
tgttgacatt ggccagggcc attgcaaaca agctggacca aaaagggttca acaattttaca 240
ccagctgcat ttaacgaggg cattgatctt atcctccgtg atgggtcacct caattgtcat 300
gcagagttag ggctgagtag atgcaggcga gctcggagac agaggacaat ggtatgggca 360
agttagtggg ctggcgctgc cggacgtggt ggtagttgcc ggatgaagtg agggcctcat 420
cccaacgtgg ccttagcttc cttggaagga ccaagcactg tgggtggcagc tgaggaaagg 480
acgatctgct tactttctgt gatctacat ctaactagtc ta 522

```

&lt;210&gt; 234

&lt;211&gt; 398

&lt;212&gt; DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:817330.1:2000SEP08

<400> 234

```
cggtctgagc cccggtgctc cctctctgca gatctccctt cgcagccgcc gaggaacccc 60
ttcacagcgc ggtatgaatg gctgcacagg agtgccccag tcgcggttga ggcaccttac 120
cccagttctt ggggcagcct atggacttga tgactgctct gcagctcgtg atgaagaagt 180
caagtgtctc tgacggtctt gtgaagggtc tccgtgtagg ctgccaaagg cttatcgagaa 240
gcacgccgct cagtcttttg gtgcttgctg aggactgtga ccagccagat tacgtcaagt 300
tggtgaaggc actctgcgct gagcacaatg ttcacctggt cactgttcct gccgactaag 360
actcttgccg agtggggccg gacttttgaca agattgac 398
```

<210> 235

<211> 687

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:817845.1:2000SEP08

<400> 235

```
ctcagagcag ccgcatcgaa ctctacctgc ctttctctcg tccccctggcg gcgtccgchg 60
caggatggcg gaatctacgg cgaggacggg gaaggatgtc aaccgcgacg agttcgtcaa 120
ggcctactcc gccacacctc agcgtcccg caagatggag ctccctgagt gggttgacat 180
tgtgaagact gcgaggttca aggagcttcc tccttatgat cctgactggt actacatcag 240
ggctgcattc attgcaagga agatctacct aaggcaaggc atcgggtgtg gtggctttca 300
gaagatctac ggtggccgccc agaggaaatg ctcccgcgcc ccacacttct gcaagagcag 360
tggtgctggt gcacgcaaca tcctgcagca gctgcagatt atgggcatca tcgatgtcga 420
tcccaagggg ggacgcctca tcaccaacca ggggaaggcg gatctggacc aagtggctgg 480
aagggtcgct gttgaagcgt gagcagtgct atctttgggt tcaatgggtg catgtttgat 540
ggttgagaag actgcgctct atttgctctt ttgggttagg attttcgagt tcagaggact 600
ttaagggtta ttgcatctag ctatggcgcc acactatggt ccccgagtac taatgtgata 660
tcaagactca attctcgaat ctagata 687
```

<210> 236

<211> 406

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:460809.1:2000SEP08

<400> 236

```
ctgcaaagaa ggggtggcgag aagaaagggt gttccgccat tagtgaggcg gttaccccag 60
aatacaccag gcgcatgggt ggagtgggct tcaagaagtg cccctcgggc acacagagag 120
atccagaaat tcgccatgaa ggagatgggg actccaaatt tgcacattga tgtgaggctc 180
aacaagctc tctgggccc aaagaataagg aatgtcccat accatatcca tatgaagttg 240
cccagaaaac ttaatgagga tgaagattca ccagacaagc tctatgcttt ggttcctaca 300
tatactggt accactttca caaatctata gacaggcaat gtggaagaga gctaaccact 360
gatggttcaa tacattaagt aaaattatth ttaaaaaaga aattta 406
```

<210> 237

<211> 626

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:815874.1:2000SEP08

<400> 237

```

tagttctaga tggcgagcg ggccgccctt gtgtgtgttg aggaagacgc ggtcgtatgg 60
gctgaggatt ttttggcccg gcacgctcct cgctcctgag ctcaccgcgt gttcgttttc 120
gccgaggaac agagtcggtc aggaagcccg cgcgcaacta tgccatggct ttttaaggatt 180
accggatata ttactacccg tggagccgga gatagcaatt cactgaattc gaatcactct 240
ctatgagcca cagcataaaa tccctggaga aggtgtgtgc tgacttgatc agaggagcaa 300
agaaaaagaa tctcaaagtg aaaggaccag ttcaaatgcc tactaagact ttgagaatcg 360
ctacaagaaa aactcctttt ggtgacggtt ctaagacatg ggatcatttc catatgagaa 420
tcacaaaca actcattgac ttgcacatcc tttctgagat tgtaagcag attagcttcc 480
atcagtactg agccaggagt tgaggtggaa gtcaccattg gagatgctta agtcaactat 540
tttaataaat tgattactgg ttgttaaaaa tatatatatt agagtactca gtctaattaa 600
acatgtgaat atattcccag gattttg 626

```

<210> 238

<211> 918

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:255713.1:2000SEP08

<400> 238

```

ctgggattgc aggagtgagc cattatgcc ggccaagtca gggcttttat agctaagatg 60
cctactacct aaaacatgta ggcaacttgg cctcctgcaa agaaggatgg tgagaagaat 120
ggccattatg ccatcaacta ggtggtgacc caagagtaca ccattaatat tcacaagtgt 180
atccattgag ctgggcttca agaagtgtgc cccttggtca ctcaaaaagt ccagaactt 240
tgccatgaag gagatagaaa actccagata tgtgcattga taccaggctc acaaaagctg 300
tctggaccaa aggaataagg aacgttccat actgtgtcca tgtgcagtgg atggtccaga 360
aaacataatg aggttagaaa ttcaccaaat aagctctata ctttggttac ctgtgtaccc 420
attaccactt tcagaaactg acagtcagac attgtggatg agaacaaact gttgatgtca 480
aataaaatta taaaactcaa aaatgaacaa atgaatgaat aaataaataa ctgtccaagt 540
ggtaacagag taggcagctg gcttgtgctg agtggaggaa cttctgtctg aatgctacaa 600
aggggagtta ttgataagga ggccaggagc ccctcagaac cttataacgt tcagggtcta 660
ggggcagggg gaattaagtg aaggtctaca tatggagtgt tgggaacctt atcctctctt 720
tccgtgctct gatctcagaa tgctggcagc caggcccaaa aaaaacacct ccagatttca 780
gcatttggag gtcacccaat aggccagggt cagcattagc ttgacctcct atacagtga 840
gtccaatgtt caataaacct taactacaca gagcatttaa tcagctttta ggacttcatt 900
cttaaaatat tctcagac 918

```

<210> 239

<211> 946

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:035973.1:2000SEP08

<400> 239

```

tatagggtaa acatccaatt aaaaggcaga tattggctat ccttggggag gccagcaaat 60
ttccggataa cggaagctg gtgggcagga aaaaggcatg caatggccct catgagccag 120
aagaatagct gctgactgat gcctggggga ccaagcccat aattgtggag ctggagccgc 180
actgtgtccc ttgttcagtc ctgatctttg gggagcctgc acagttctga cctcagttaa 240
gagctaaggg aagagatttg ggaatgtttg gccctgtgta acagcaaaaag gattcatttg 300
tgagggtgtc tggcaggat cctgaaagat aataagatga agggacatc gccgtttgga 360
aagtgtcgtg atatgataca caagttgtgc tgcctctgtg gctctaaggc ctaccacctt 420
cagaagttga cctgtggcaa atgtggctac cccgccaagc gcatagtaga aagtatatac 480
ggcagtgcca aggcagaat agacttatat acctctggga ctggtcgaat gaggcaccta 540
aaaaattgta gatccgcaga ttcaggcaca gattccgtga aggaacacca cctaaatccc 600
aagtagggca tgctgctgca gcatccagtt catcttaaga atttcaacga tgagtaatgc 660

```



aatagatggt ccggttttaa aataaaaaat tttatgaata cttaaagtttg aactttacat 720  
aattttataca tgtcaccata taataactctt ttgggtttttt tcagtcctta aaaagaatgt 780  
aaaaaccatt ttagcctgca ggctgtacaa ggacagatga cacgctgggtg ttggttctaa 840  
ggccaggcta tagtagtttg ctaatccctg ctgtagagga tagcatcggc tactgatgaa 900  
cattacgggtg ttacaatgat tattcttaga atttttaaga acacccc 946

<210> 240

<211> 618

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1138110.1:2000SEP08

<400> 240

aaaaaataaa aaataaaaaat aaaatttttaa tgtaagaaaa ttgaaggaaa aaatggactt 60  
tgtgttatgt cttcccttac tttgttctgt ctgtatattc ctatattgat tcggtagtag 120  
caatgcataa ttttcccaca aacttcatec ttacatgtag agaaagtgtt gataaaatat 180  
tctgcaataa agtgcctctt gcaaatatgt acttcatttt cactgtttat tcaattttct 240  
taataccata taagttctta caggagagct tcagattctc catacaaaat ggttaatcta 300  
aaactatgct ggatagagta aaacccttgc tgagcttctt agaaaaaaaa aaaagatgct 360  
ttatgtgtca aatataggct atgttaaate atagcaaagc ctttggcctc atggcactcc 420  
cgaaagcaag atgggggtcac cagcagccct actgggagcc acccacggaa gctgggccag 480  
gttctcactc tggtcacatc tgcccaaacc agcatagtct gatctggaaa cacggcctca 540  
atatgtgccg ccggtgttcc catcagtatg caagggacat aggtttcatt gagtgggact 600  
aagtgatctt ccttgaat 618

<210> 241

<211> 720

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:2049074.1:2000SEP08

<400> 241

gttctcctgt acaggaaggc agccatggcg cccagccgga attggcatga tcctgaagcc 60  
ccacttccaa aaggaatggc aagcagggag tgggacacgt ggttcaacca gccggcccgc 120  
aagatccgca gacgcaaggc ccggcagggc aaagcgcgcc gcacgcccc tcgccccgcg 180  
ttccgggtccc atcaggccca tcgtgaggtg ccctacagtt agataccaca tccaaggtcc 240  
gggctggcag gggcttcagc ctggaggagc tcagggtggc tggtatccac aagaaaatgg 300  
cacgcaccat cggcatctcc gtggacccaa ggaggcgaaa caaatccacg gagtacttgc 360  
aggccaacgt gcagcgcctg aaggagtacc gctccaagct catacttttc caccaggaag 420  
tccttctgct cctgaagtaa gggagacagt tctgctgaag aacttaaatt ggcaacgcag 480  
tctaactatg gtacctgtga tgctctatcc ggaatgtgta caaaaaggag aaggccagag 540  
ccatcacgga agaggagaag aactttaagg ctttcgccag ccttcgcatg gcccgagcca 600  
atgcccggct cttcggcatc cgagcaaaga gggcgaaaga agccgcagag caagacgttg 660  
agacgaagaa aatatgcgcg gtggagagtg aataaatttc cataagcaaa gaaaaaagg 720

<210> 242

<211> 670

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1092460.1:2000SEP08

<400> 242

gcggcgccgc cgcggggctg cagtgggcgt cgcgccggcg agaccatggc tggaggcaaa 60  
gctggaaaag acagcgggaa ggccaaggct aaggcagtg ctgcgtcaca gagagctggg 120

```

ctccagtttc ctgtgggccc catccacaga cacttgaaga ctgcgaccac aagccatgga 180
cgggtggggcg ccaactgctgc tgtgtacagt gctgcaattc tggagtagcct cacagctgag 240
gtgttggaggt tggcaggtaa tgcttctaaa gatctcaaag taaaacgcat caccacacgt 300
cacttacagc ttgcaatccg aggtgacgaa gagttggatt cactgatcaa ggctaccata 360
gccggggggcg gtgtgattcc gcacatccac aagtctctga tcggaaagaa ggggcagcag 420
aaaactgctt aggacgtgag tcaactgacct gctgtcctca ctgtgtgtga ctgggcagag 480
ggtaccagtc ggtgtgtggg aaagcataga caattgctgt agacagaaga cacatttgta 540
tgtttttaga ctctcgaagt ttgattaaac tcctttccct gtggtcagtc atgtgtttga 600
ctggccaggt ctctcccttg tgttttatac cacaaataga attgctcaac atttttacaa 660
caggaaaaaa

```

&lt;210&gt; 243

&lt;211&gt; 2552

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:399421.1:2000SEP08

&lt;400&gt; 243

```

caaaaacgaa gcccacagct gtttcagtat gtttaagtgg ctctgaaaag agcctttggg 60
ttagaaatga cgttgtacct tagcaggctt acttggagct ggtgtgcttg gtgacagcct 120
tggtgccctc agacacagca tgtttggccg gctctcccgg cagtagcaag cgcactgctg 180
tctgaatctc tctggaagaa atggtggagc gtttgctgta gtgagcctaa cgtgatgcct 240
cgctcgctat agctcaaaag atatcagtga cgaaggaatt cataatgctc atggccttcg 300
aagagatgcc tagtgtccgg atggacctgc tttagcactt tgtagatgta aatagaataa 360
ctctccttgc gggctcctta tgcgcttatt agccttcctt tatctggggg cttgacgaca 420
gctttcttga agcccttctt gggaaaatggt agcaccttta gaatgacacc tccggcatag 480
ctaccacatc acttctgcag cggtacacgg atcttttcca gggttacaca tgcactttcc 540
agtgtccaac agctcttcag tattaacacag ctccaaatga cctcagctcc aagtctcgcc 600
aagtaacaa acagaactgg aggatgaaag atgctggtat ttatatggca ctggccctca 660
aactgcctgc gttctggtat ccaaaatctt gatttggaca catgtggaat gtgtgctggt 720
gaattgcaat atgatgtgct gaattctatg tactgacttt cttgatttgt gtgtatgact 780
gtgccatcat gccaatcatg atatgcaaat gccaacctca tgctcatgact ataaataggg 840
ccagttagca gccttagaaa tgtggttctc ggtgtgtctg gcatttgatg cgaggagat 900
gtctgggacg aggggaagca ggggagggaa aaaggccggc gcccaagttc taaagtctcg 960
ctcttctaga gcgggtttgc agtttcccg taggcccggat ccacgtctcg cttcgtaagg 1020
gaaactatgc agagcggata ggagcaggcg caccagtgtg tttggcggca gtgttggagt 1080
atctcacagc acatactcct tgtaggaagc aggcaatgcy tctcgcgata acaaaaaaac 1140
ttcgcatcat tgcccccca cctgcagctg gcgactccgc aatgatgagg aactcacata 1200
agcttttggg cgccgtgacc acttgccac gggcggagtc ctgcctaaca gttcaggcag 1260
gtgctgctgc ccacagaaga ctgagagtca ccaccataaa gcccgaagca agttacttaa 1320
ggttgtcaat tgataaaaag acttctgtgt cagaaaaaac aaaaaggctc ttttcagagc 1380
cacctaccgt gctcagagaa aaggctgtac gttaccaaac gttactgctg ttaaccagcc 1440
atcatcacac gcgctcgtgt cccaaatgcc cgtttaatgc gcatttcggg aatgtacggg 1500
gactgaatth tcaaggagcg tttccatcgg ccattagttt gaggaatth tggttatata 1560
tgtcactggg cgagggtggc taacgcttgt gagcagctca gtacttgca ttcagtgcga 1620
ggttggagcg ggagggtcgt tcaagacctg gggacttgac cgcttaacag taaattgaga 1680
gccctgtccc cacagaaact acgactgggt ctggtgatgc aggtcctgta gtctcattht 1740
gttcgggatg gctgcggtgt gtagggtttg ttcagtgtc cgggtaggcg aatagtttca 1800
gtgatacgta agattgtatt gctccactgc actccagtct ggggtgacatg agcgagacta 1860
tctcaaaact gagttatgta tgttacgtta tggttaacggg ttacacatgg aagtgaataa 1920
tgcattacaa gcaggagagg tgtaatcttc ctgtttaaat ccaggtaaca tcagtctggc 1980
tggttaaagg acaatttaag gggatattcc tgtttcaaat tatggtgggt gttcacgaga 2040
atactccata attaaatagt gggcatagtg aagtgaaggt aatacttaaa atattgggtg 2100
aagcataata ataatttggg taacctcgtg ggtgttcacg tgactcttgg ttaaaggatt 2160
aataggatth attttcttaa aatagtgaag taaaataata gtgcaaaaag aaagcctcta 2220
cgccactaat gtgagtgggt agggagtaact atgtgttatg agactcacgg cacggagatg 2280
ggagactaga gaggccgaga acaataaatt tgagaagttt agagaaatth ggagagaggt 2340
aataatgctt cacaaaagtt ttgagactgt tgagaaagag ggcacaaaat aagtttgttg 2400
agactctgtg ttgcagtttg tgctgttaaac accacagaga aagtggaaag ccaaacccat 2460
tcagtacgct gcccaataac ttaatttaaa attattgtac gatattccac agttgacctt 2520

```

acctccaaaa aaaaaaaaaa tcatttggtta cc

2552

&lt;210&gt; 244

&lt;211&gt; 1299

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:816655.2:2000SEP08

&lt;400&gt; 244

```
agaagttgga gcagtggtgaa ggaagaggtg agaacaaccc cgggactgat caaagcccac 60
atggccactg cattcccgcg gtccagggcc ttacatcctg ccaccgtcgc tggcaccacc 120
atggcccaag gaggaagggc taaaggggat gctaaaggag acaaagccaa ggtgaaggac 180
gaaccataga gaagatccac aaggttgtct gctaaacctg ctctccaaa gccacagccc 240
aagcctaaaa aggccctgc aaagaaggga gagaagatac ccaaaggga aaagggaaaa 300
gctgatgctg gcaaggaggg gaataaccct gcagaaaatg gagatgccaa aacagaccag 360
gccccttgga gctgaagggt ctggagttgc caagtgaagt gttgtgcatt ttgaataact 420
ggtgtacttc ggtggctgt acaatttcaa atactaatgt ttatcaagt ttataaaaa 480
atgcagaaat ttgtttaact ttttttttta agcccatgtt gttaattcac agaacttc 540
attgtgttt tatgggaag ggcgtatgt cactagtaga atgtctcaa gagctggatt 600
gatgtggaga aaacacctt ccttctagt tttgagagac ttccctctgg cctcccagga 660
ggaggggatt cccgtggact ttgacacaca gtggccacct ttggcacaaa agtccttgtg 720
gtatagaaaa acaaatttgt tgtttactgt cctcttctcc cttccatct ttcagcatag 780
acttaactcc cttaagcccc agacttctgg ttgagaacct gacccctag ttcatttgg 840
taccaggtgt gtgtcaggcc aatctggaac tttcccagtg catgccactg agagtggcac 900
cgtgtgtcaa aatgagcagt gggttccatt tctaggattg tgggatcttt cagataaatt 960
cttgccattt ttctttcac cttctgaaa ggtcagcggg ccggctttgt tgaacaagg 1020
ttgttaaacc aaccatgctt aaatgtgaaa ttgtcaaccc cctcactcta gaactttccc 1080
tggtcaagaa catcaggatt gaagacttcg attgggtttt aaagtgggct ttctggatat 1140
ttggggaag tccacttgta agaaaggag ttttgaaagt tgggttggtat actggttaac 1200
gattgtctgc ccatggtcct gctgtgaaat accatgattg tttatggtaa gtatctttaa 1260
taaagctgga tacagttggc tggaaaaaaa aaaaaaagg 1299
```

&lt;210&gt; 245

&lt;211&gt; 677

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:414732.1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 45, 618

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 245

```
agaggagcga aaagcatggt ccggggggcg gcggagcgct ggtangcgct cgggcccgtg 60
ttgtggggct gcgcgctggc gctgcagggc gggatgctgt acccaagaga gaggccgtcg 120
cgggagcgga agaagctgga ccgcctctgg agcttccgcg ccgacttctt cgacaaactg 180
tgccttggct tctaggagta gtggtatcgg cgtctgctgc gagagtggg ctccaccatg 240
gacataccgg ttccctccag cttcaacgac gttggccagg actggcggct gcggcatttt 300
gtagaccaga tgtgttacga acgggaggtg accttctctg agcaatggac ccaggacctg 360
cacacaagag tggtagtgag gattgtcagt gccactcct atgccatcgt gtgggtgaat 420
ggggtcgacg cgtagagca tgaggatct acctcccctt tgacaccgac atcagtagcc 480
tgttccagggt gggggccctg cctcccgc tccgcatcac tatcaccatc ggcaacatgc 540
tcattctctc caccctgcc aagggagca tctctgacat ggcggacacc tccacgtggg 600
taccatcctg cttccacngc agacacccac cttcctgtcc caccocgtgg ggcattacat 660
tagggtaatt acattag
```

<210> 246  
<211> 649  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:1140250.1:2000SEP08

<400> 246  
gtgactggag gtgtatcctt gtcctcgtct gaatcacgct gatgtggccc caacccacc 60  
tccctcccca ccccatgatg tcagaaaaaa ccagacagaa caaattggct gaggccaaga 120  
aaaagtttac agactatcgt cagtggaaaca ttgctgggtg tggtagcgga gcaactgaca 180  
ccaaaaagaa gaaaaataat catggcacta accctgagac aaccacttcg gggggctgcc 240  
actgcctga ggatacacaa cagaaccgag cgcagctgaa agaagtaacg tgatttggtt 300  
gctcacgaca tgactgctgg gtttggaggg cactcagatg tagaggcccg tctcatctcg 360  
cctactccca gcttggggaa gaaggctcac ccccaagatt ccaccccatc cccacagggg 420  
ccctgataac ctggtcccat gggcgccgct gtcctggggc attggtgccca ttctgggggc 480  
atgtctcttg ctgtggatct ctgctcccc ctagtaagag ctctgtcttc ctctttctat 540  
aggaaaagaa ggcaagccac caacatcagg aaagccctaa gcgggagata gaggccagg 600  
atcatacaat acgaattctc acgtgtcaga aaactgactg gagacaggc 649

<210> 247  
<211> 1016  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:174022.1:2000SEP08

<400> 247  
gctctccgcc gcagcttcag cacttcagcc cagaacaatg ctaaagtagc tgtcctggga 60  
gcttctgggg gcattgggca acccctttca ctcctcctga agaacagtgc ccttcgtgag 120  
cctgtgacc ctctatgaca tcgctcacac acctgggtgt gcagcagatc tgagtacat 180  
cgagaccaga gcaaatgtga aaggctacct cgggccggag cagctgccgg actgcctaaa 240  
agggtgtgat gtgggtgtca tcccagctgg agtgcccagg aagccaggaa tgacacgaga 300  
tgacctgttc aacaccaatg ctaccattgt ggccacattg acggctgcct gtgcccagca 360  
ctgtcctgaa gccatgattt ggattatttt ccaacccagt taactccacc atccccatca 420  
cagccgaagt tttcaagaag catggcgtat acaaccccaa caagatattc ggtgtgacaa 480  
cccttgacat ccgtcagaag caaacacatt tgtggcagag ctaaagggtt aggaccagc 540  
tcgagtcaac tgttgccttg tcattgggtg gccacgccgg gaagaacgat catccccct 600  
gatctctcag tgtaccccca aggcttgact ttccccatag gaccagctgg ccactactca 660  
ccgggaggat ccaggaggct ggctactgaa gtcgtgaagg ccaaggctgg agctaggctc 720  
tgccactctg tccatggctt atgctggagc ccgctttgtc ttctccctgg ttggacgcca 780  
tgaatgggaa ggaaggagtc atcgagtgtc cttttgttca gtccataaga gacagagtgc 840  
acttatttct ctacgacct tgctgttggg gttattatag gcctggtagt atagtatacc 900  
ttaggctatt ggcaaaatct actcctgttg gaggtagaaa tatgattgcc gaggccatac 960  
cctgagctga aagcctccat ctagaaagga gaggactttg tgaagaacat gaagtg 1016

<210> 248  
<211> 890  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:002811.1:2000SEP08

<400> 248  
aaaaaaagaa gaagagggtta gacttaactg tttccaagat ttcattccagt tgtgaaatcc 60  
tgttatctga actgccaggc ctacaagttc aattcaaaaag atggaatgta gcttctctgca 120  
gtcttttaaat gaacaaaaca aacagcacia gctacactgc gaagatgagt ctgcactaca 180

```

tatcttttctt tttgccccca ggaacctggc ttaccagtat atcacccag cgttgacaca 240
ctggtgagtg acgacagtgc gctttttgca aatcctgtgc actatgactt ttttccttca 300
agaatagatt tgactgggtct gaaattctat tttcatattt ggctgtgctc tagtaataac 360
ctgcctcact ggacatgtgg aattgtcctg cggttcaggg aacagggaga gaacagggct 420
ggctctcctc tgctctctggg ggcccagcct cccaaccggc caagtccagc tctctcctga 480
tgagcttcca gtagctcact ctctcctgct gctaactctt gtttcctctt ctaaccatga 540
ccctgttggt ccctgaatgc gcgatgtttt atcacactgt agtttatgac atttggaagt 600
acctgaggat ttaatgaaaa ggaaaacttt ccgccatttt ccttaatgtc agttttagc 660
tttttcttct gctgcccagg taagtacttc actttcaaag agaagataca atttgcaaa 720
attcccatca ccctcctgta ggaaacaaag tttttaagga ccctctctta agtcaaatg 780
acagtgaatt tagaattgtt atacatataa aatggaaaaa ttaggcatta ccttaagaaa 840
cattgtaatc atgttgtaat tgaatacatt gtaatatatt gaatatcata 890

```

&lt;210&gt; 249

&lt;211&gt; 751

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:414732.2:2000SEP08

&lt;400&gt; 249

```

gagagaggag cgaagagcat ggtccggggg gcggcggagc cctggtaggc gctcggggccg 60
ttgttgtggg gctgcgcgct ggccgtgcag ggccggatgc tgtacccaag agagagggccg 120
tcgcggggagc ggaagaagct ggaccgcctc tggagcttcc gcgccgactt ctccgacaaa 180
ctgtgccttg gcttctagga gtagtggtat cggcgtctgc tgcgagagtc gggctccacc 240
atggacatac cggttccctc cagcttcaac gacgttggcc aggactggcg gctgcggcat 300
ttttagtagc agatgtggta cgaacgggag gtgaccttcc tggagcaatg gaccaggac 360
ctgcacacaa gagtggtagt gaggattgtc agtggccact cctatgccat cgtgtgggtg 420
aatgggggtcg acgcgctaga gcatgaggga tctacctccc ctttgacacc gacatcagta 480
gctgttcca ggtggggccc ctgccctccc gcctccgcat cactatcacc atcggcaaca 540
tgctcatctc ctccaccctg ccaccaggga gcatcctcga catggccgac acctccacgt 600
gggtaccatc ctgcttccac cgcagacacc caccttctg tcccaccccg tggggcatta 660
cattagggtg attacattag gagaacttgt tgtttaaaag agtgtggcac cttcatccac 720
cttttctctt tcttgccatg tgacacacca g 751

```

&lt;210&gt; 250

&lt;211&gt; 474

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1019920.1:2000SEP08

&lt;400&gt; 250

```

gtccttcgga gcggaaggaa tatggcgctc gctactcgcg ttatccaaaa gcttcggaac 60
tgggcgtctg ggcaagacct gcaggcgaag ctgcagctgc gctaccagga gatcgccaag 120
cggacccagc cacctccgaa gctccctgtg ggccccagtc acaagctgtc caacaattac 180
tactgtactc gtgatggcgg ccgggaagtt gtgcctccct caatcatcat gtccacaaa 240
aaggccctgg tgtcgggcaa gacagctgag agttctgctg tggcagccac taagagggca 300
gtgacacctg ctccctccat gaagaggtgg gagctgtcaa gggaccagcc gtacctgtga 360
ccctgagttg gctaccttgc tatgtttcca gggccacatg actgcttttc ctcttggat 420
tccttctggg gagagtgtga cctaatttat aacaaatata ttaagtacca catt 474

```

&lt;210&gt; 251

&lt;211&gt; 457

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: LI:1038336.1:2000SEP08

<400> 251

```
aatttttaaat atttgaagat gttgaattaa ctgagtgaga cttaaaaaaa aatttttcagg 60
aaattttcaga atgctctatg tctcgttgc cattggcggt gccttttctca agaatgcctg 120
gcctaaggag cgggtgctgg tectgttctt caccatcggg ggcttagcta taaatctgac 180
cccagtctgc ccctatacca tgtactccac caggatcaac taggccacgc ccacaacaa 240
cccagtgcct ctcgagatg atgggaacat gctggacatg cccagccacc cccaaggccc 300
aagcatggag tggttgaaga aactgcaagc acttccattg acatggagaa gaccatgtcc 360
cctgtggccc ccaataaaaa tgtgaccccc cccgcctccg aaataataa atcattcagt 420
ctctgaaaag tgtttggtta aatgtgctag acttata 457
```

<210> 252

<211> 3993

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1177772.11:2000SEP08

<220>

<221> unsure

<222> 2234

<223> a, t, c, g, or other

<400> 252

```
gaaaggtaaa ccagtaaaat aacattgtgc ttggtgctga acctatgcta gtattggcat 60
taacactgac ctttatttaa ggttctaatt tgttcatggt gggcaacttag aacatcggtt 120
tggtgttttt tttttgtgag attctggaaa cattccaatt ttaccttttc cccttgactc 180
cagactttttt aacactggtc tgctactggt aagttgtatg ccattttggt aggccttctc 240
aagtgggaat caggaacgct gctgtgctct agagatgttt tgttcttctt gtagggctga 300
agcagtgcct actcgataga atcagtcac atgcaaataa aagccacctg agtcaaaggc 360
aaagccagag tgagcgttgg agcaaagaag ttaaaagcat attggcagag gaagagccct 420
ggcattccag caggagctaa caggaaaaag aaaatcaatg gcagtagccc tgacacagcc 480
acttctgggt gttaccactc acctggggat tcagcaacag gtatctacgg ggagggccgt 540
gcctcctcta ctaccctgga ggatctggag agccagtacc aagaactagc agtggccctg 600
gattcaagct ccgcaataat cagtcaactc actgaaaaca tcaattcact ggttcgcaca 660
tctaaggagg agaagaagca tgagatacat ctggtacaga agcttgggag gagcttggtc 720
aaactcaaaa accagacgga acaacgagaa caagagcaca ctgcagtttg agcagcaagt 780
aaaggagctg aagaagtcag gtggagctga agagccaaga ggctccgagt ctgcagcagc 840
agccagacca gtgactggag cccagctccc acaaggagct tggatgtgag gacaagcagg 900
gtggactccc caggagcacc caggcttgag tggagaagct gttggtacag gagaggcggc 960
aggaggagca ggagaggctg catgccattc ttttcgggct gcggagaaca gggagctaaa 1020
catcaccatc atctaagagc ggttcaagga attgaaaaaa aaaaaaataa catttaaggg 1080
gttaatatcc tacacaattc atttaactta ttgaatgtt agagccactt atgtttattt 1140
gtgtttctaa tttatagttt aaattttatt gtgtttctaa tttatagttt aaattttatt 1200
ggtgtttcta atttatagtt taaattttatt tgtgttctaa tttataattt aaattttatt 1260
gtaaaaagtt aaatgagagt ggtgtttctc tcatgttcac tctggcatct tttagcattt 1320
ttttaatttg atattatagg acgttagcat gcatatcgag tttgccctta tgtggtggga 1380
gtcaaacaca caaagacca ctgtatgcac acaactgttc ttgctggtt ggataggct 1440
gccatgcttt tttaatgtta gtacagcctg ggtgacagag caagactcca tctcaaaaaa 1500
acaaaacaaa aaaaacagaa tgaaaaaggc aaaccagtaa aataacattg 1560
tgcttggtgc tgaatttatg ctaggatagg cgtaaacagt gacctttatt taagggttta 1620
atttggtcat gttgggcact tagaacattg gtttgtgtt ttttttgtga gattctggaa 1680
acgttccaat tttacctttt ccccttgact ccagactttt taacactggt ctgctactgt 1740
taagttatat gccatttggg taggccttct caagtgggaa ccaggaacgc tgcgtgttct 1800
tagagatgtt ttgttcttcc tgtagggtg aagcagtgcc tactcgatag aatcagtcac 1860
catgcaataa aaagccacct gagtcaaagg caaagccaga gtgcagcttg gagcaaagaa 1920
gttaaaagca tattggcaga ggaagagcac tggcattcca gcaggagcta acaggaaaaa 1980
gaaaatcaat ggcagtagcc ctgacacatt catttctggt ggttaccact cacctgggga 2040
tgtgagcttc ggcgggccag gctcctgggg acagggagcc caaggggcag tagagggtaa 2100
ttgttaagat tgtagatgga ctgttgggta ctggttaaga attctggatt tgaatcctgc 2160
```

```

ctctccgtct gctaagaatt gattagggat tgattagcat atgatttagg gcaatttgct 2220
tgagggtcttt ggcncctctct tttcacatct gtataataga ggtgggtatct tttgacttcc 2280
atthtgaag tttaaatgag attcggttatt gttgctttat gtgaatcctt agtacatggc 2340
ctgctgcaaa caccaggac accgagggaaa tggctgttgc tgtttgattt tctcatccc 2400
cagttctcaag gggaagccag gccaatgaga agagccactt gccatcaggc tgtcccttta 2460
ggagtcaactg aaagggtccc aggggtgggat ggtggggaga taagaaccac gagagaagtt 2520
ggcacaagag agttatggga aaaagggttcc aagataggca gaaagaagc ttttgccagt 2580
tgatggggga agaaaggaag tcagagggct tagacagtga ggggggacag aacatctcca 2640
tgtgcaactct catctcttgg agtcagcaac aggtatctac ggggagggcc gtgcatctcc 2700
tactaccctg caggatctgg aggtaaagg ccctgggccc aggtgcagtg accctgcagg 2760
ccagccctcc aacctctccc cagcagggg gcttgttgc cctctgccag ctgaggcagc 2820
ccacacaccc ccaccagccc taatgattat tctctctacc cctccccaca atcttctccc 2880
aactccttct ctctgcatgc acctcagagc cagtaccaag aactagcagt agccctggat 2940
tcaagctccg caataatcag tcaactcact gaaaacatca attcactggt tcgcacatct 3000
aaggaggaga agaagcatga gatacatctg gtacagaagc ttgggaggag cttgttcaaa 3060
ctcaaaaacc agacggctga acccctggcc ccacagccc cagcagggcc atctaagggt 3120
gagcagctac aagatgagac caaccaccta aggaaggagc tagagagtgt gggagacag 3180
ctccaggctg aggtggaaaa caatcagatg ttgagtctcc tgaacaggag acaggaggag 3240
aggctacgtg aacaggaggga gaggtacgtg gaacaggagg agaggctgtg tgaacaggag 3300
gagaggctgt gtgaacagga ggagaggcta cgtgaacagg aggagaggct gtgtgaacag 3360
gagaagctgc cagggcagga gaggtctgct gaagagggtg agaagctgtt agaacaggag 3420
aggcggcagg aggcagagga gaggtctgct gagagggaga ggctgctgga cgaggaggag 3480
gagctcctgg agcaggagag gcttcggcaa caggatgaga ggctgtggca gcaggagact 3540
ctcggggagc tggagaggct ggggagctg gagaggctgc gggagctgga gaggatgctg 3600
gagctggggt gggagggcct gtacgagcag cgggcccagc cagcagcgg cttcaggag 3660
ctgaacaacg agaacaagag cacactgcag ttggagcagc aagtaaagga gctggagaag 3720
tcgggtggag ctgaagagcc aagaggctcc gagtctgcag cagcagccag accagtacct 3780
ggagccccat tcccacaagg agcttggatg tgcggacaag cagggtggac tcccaggag 3840
caccaggctg tagtggaga agctgttggg acaggagagg cggcaggagg agcaggagag 3900
gctgcatgcc attcttttcg ggctgccgag aacaggggagc taaacatcac catcatctaa 3960
gagcgggtca agaaattgaa aaaaaaaaaa agg 3993

```

&lt;210&gt; 253

&lt;211&gt; 710

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:205642.2:2000SEP08

&lt;400&gt; 253

```

cgcgacggaa accccaaacc caacctcatc tcttttctcg gcctctaccg cccctctgat 60
cggtcggccg gtcagctatg acgcccggaa gctcggccgc cggatccggc gtcgtcgttc 120
ctcggaaactt caggctccta gaagagcttg aacgtggaga gaagggcatt ggagatggga 180
cagtgaagta tggaatggat gacgcagatg acatctacat gcgatcatgg actggcacta 240
ttattggccc tcataatacc gtccatgagg gtcgcatcta ccagctgaag ttgttctgcg 300
acaaggacta ccctgagaag ccaccatcag ttcgatttca ttcaagaata aacttaacat 360
gcgttaatca tgaaactgga gtggttgacc cgaagaagtt cagcgttctg ggtaactggc 420
agcgtgatta ctcaatggaa tacatcctaa cccatctcaa gaaagagatg acatcgccac 480
agaaccgcaa gctagtccag cctccggaag ggacattctt ctaaagatgg cagcaactca 540
tgtccatcca tgggtactcg gaaacagctt ccgcacctgg ttcctatatt ttatatattt 600
agtctgatct cccagttatc tgtaaacacc cgagatcgag tgcaagcaga cttgggttgat 660
ttacaatttg cccgggtggt gctaatagata tgggatgaaa cggttgccgg 710

```

&lt;210&gt; 254

&lt;211&gt; 613

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:449685.1:2000SEP08

<400> 254  
tttcggctcg agagccattt gaacccccca cgtctcctcg cctcggtccc catttgctgc 60  
gtctccttgc ccgcgcgcgc gctcgcggtt tgccccctgt ctccgcaccg cgcgctgttg 120  
tccatccttg tgcagcaaaa atgacgggtg gctacattgt tggctcactg gtcggatcct 180  
ttgccattgc atacctgtgt gacacattta tctctgacaa gaaggcattt ggaggtagca 240  
cccccaagac tgtttctgag aaggagtggg ggcaagccac agacaccaag ttccaggcct 300  
ggcctcgcac tgctgggcca ccggttgtca tgaaccccat cagccgccag aacttcacgc 360  
tcaagtccac tgagtagggt tgggctggag aagctgggtc tgcgacttgg tgccgccatg 420  
aagccagagt tccttttcta aggtcttaac caattgtaac aaaacgaaat acagtcagga 480  
tttttttggg atttttactt gatctgtgag cgaacactgg gtctttcacc ttgtcatcca 540  
atgggttctg ctcccttggtg aaatctccca ccagccagta aagaactggg ttgttttcat 600  
aatatatgct ttt 613

<210> 255  
<211> 634  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:453922.1:2000SEP08

<400> 255  
cgtcagcccc tctcgttttt catccgcctt cgcctccaac cgcgtgcgct ccacgcctcc 60  
tccaggaaaag cgagagcttc tgcataagaa caacaaatca aaagcgtgat cagctcggtg 120  
ccaacaaaac ctcaacaacc aagtttcatg tctgatctcg acgtccagct tccatctgcc 180  
tttgatccgt ttgctgaggg aaatgctgag gactctgggt ctggtcctgg aacgaaggat 240  
tatgtgcatg tgcgcatcca gcagcgcaac ggcaagaaaga gtctgactac agtccagggt 300  
ctgaagaagg agttcagcta taacaagatc ctcaaggatc tgaagaagga attctgctgc 360  
aatggtactg tagttcagga cccagagcta ggccaggcta ttcagctcca aggtgaccag 420  
cgcaagaatg ttgctacttt cctagttcag gctgggattg cgaagaaaga gaacatcaag 480  
attcacgggt tctaagggac ctgtaaatgc ttgtgcccta tattgtgtgc ctcaacatat 540  
tggggagcct gaagcatcga cagttgctag tcattgctta cttatataag aacataagta 600  
gtatttgcta gttgtcaagt gtgccttgct tgat 634

<210> 256  
<211> 550  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:476342.3:2000SEP08

<400> 256  
gattgatcgg tcacctgacc tcatagataa ctctaggggc gacggacgga cgcgcgtctc 60  
cgggtcccgtc cgtaatatga ctgatccgat ccacgcgggc cggcgatgga gctcatcaag 120  
tccagggcga ccgtgtgcgc gctcctcctg gcgctgctcc tgctctcgca ctacgacggc 180  
gggacgacga cgacgatggt ggcggaggcc cgggtgtgca tgggcaagag ccagcaccac 240  
tcgttcccct gcatctccga ccgcctctgc agcaacgagt gcgtcaagga ggacggcggg 300  
tggaccgccg gctactgcca cctccgtac tgccagggtg agaaggcgtg ctaagcaaag 360  
ctcttgaaac acccttgggt tgccagaact gaactgtggt agtactaagt aacacccttg 420  
gctagctgtg cacaacctac gtaccgtgca tgcattgtaat gtggtgtcat gtaacgtgac 480  
agcaataaat attaataaca ataataacac ggcattgtagc ctttgcattg ttccaaaaaa 540  
aaaaaaaaag 550

<210> 257  
<211> 756  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature



<223> Incyte ID No: LI:336801.1:2000SEP08

<400> 257

```
ctttgatgtc acaatctgac ctgtgggtat taatctggtc acagctgctc acagactaga 60
tggcaaggtc caggggacta gcacaggaag atgtggaaga gaagtgacca tcaacccaag 120
atcaaagcag aggatggacc tctggtgggc cagttcgagg ttctgggttc agttccagaa 180
cctgccatgc cacatcctct agagctgtca gaatttgaga gcttcccagt gtttcaggac 240
attaggcttc acatcagggg agtgggagct caattgggtca agaaagtcaa tgcctgtctt 300
cagctggaca tcacaaaaaa tgggaagacc attttgcggt ggaccattga tctgaagaat 360
ggttctgggg acatgtatcc gggacctgcc aggtcccag cagacactgt ctttacaatc 420
ccggagtctg tctttatgga gctggttttg ggcaaaatga acccgagaa ggctttcctt 480
gccggaaagt tcaaagttag tggcaagggt ctgcttagct ggaagctgga aagggttttc 540
aaagactggg ctaaaactttt aagcatgcaa aaataccaag gaagaacttc tcttcgatga 600
tatgtcgagt ggaatcatg cttaatctga agattaaagt aaaaagtatc caatattttt 660
actcaaattc ttcctattca agtttaattg attttcctgt tgatgggtaa acctccaggg 720
agggttctga agcagtaaaa agagatgtag tgcaaa 756
```

<210> 258

<211> 616

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:449685.1:2000SEP08

<400> 258

```
tttcggctcg agagccattt gaacccccca cgtctcctcg cctcgttccc catttgetgc 60
gtctccttgc ccgcgcgcgc gctcgccgtt tgcccdctgt ctccgcaccg cgcgctgttg 120
tccatccttg tgcagcaaaa atgacggctg gctacattgt tggctcactg gtcggatcct 180
ttgcctattg catacctgtg tgacacattt atctctgaca agaaggcatt tggaggtagc 240
accccaaga ctgtttctga gaaggagtgg tggcaagcca cagacaccaa gttccaggcc 300
tggcctcgca ctgctgggcc accggttgct atgaaccccc tcagccgcca gaacttcac 360
gtcaagtcca ctgagtaggt ttgggctgga gaagctgggt ctgcgacatt ggtgccgcca 420
tgaagccaga gttccttttc taaggtctta atcaattgta ataaaacgaa atacagtcag 480
gatttttttg tgatttttac ttgatctgtg agcgaaactt ggtcattca ccttgtcatc 540
caatgggttc tgcctcattg gtgaaatctc ccagccagcc agtaaagaac tgggttggtt 600
tcataatata tgcttt 616
```

<210> 259

<211> 444

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:476342.1:2000SEP08

<400> 259

```
ggacggagcg gcgtctccgg tcccgtccgt aataaactg atccgatcca cgccgggccc 60
cgatggagct catcaagtcc agggcgaccg tgtgcgcgct cctcctggcg ctgctcctgc 120
tctcacacta cgacggcggg acgacgacga cgtgggtggc ggaggcccgg gtgtgcatgg 180
gcaagagcca gcaccactcg ttccctgca tctccgaccg cctctgcagc aacgagtgcg 240
tcaaggagga cggcgggtgg accgcccgtt actgccacct ccgctactgc aggtgccaga 300
aggcgtgcta agcaaagctc ttcaaaccac cttgggttgc cagaactgaa ctctagtagt 360
actaagtaac acccttgggt agctgtgcac aacctacgta ccgtgcatgc atgtaatgtg 420
gtgtcatgta acgtgacagc aata 444
```

<210> 260

<211> 640

<212> DNA

<213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1072804.1:2000SEP08

&lt;400&gt; 260

```
ctcctaccta cgcagcagag gccagccacc tccccgtcgc tctcacgctc ttcgtgcccg 60
ccgcctctcc tccgtctctg cctctctccc tctcccctcc cgtctgacgc acatggcctc 120
catcgccgga tcttcggccc tctccttcgc caggcccgtc aaggcaatca acacaaattc 180
tctagctttt tccccggcga ggaagggcaa cacattcctc cgcctgcaac caatgcctat 240
gagatctggt tcttgcgctg ccaagaagga tacaacagac aaggtttggt agattgtgaa 300
gaagcagctt gctcttctg accatactga agtttgcggt gaatcgaaat tctctgaact 360
cgggtgcagat tcaactggaca cggttgagat tgtgatgagt ctcgaggagc acttcgatat 420
tagcgtggag gagtcacgag cacaacaat cgccacagtt gaagacgccg cggacctcat 480
cgacaagctt gttgctggca aagcgtagtt ctgtacctcc tctatcgttg tctaccact 540
ctaggtctat cccttgatct aggtttctat ggtctgaaga tttgtgattt tctccgcact 600
gttggtgcta tccccgtgct gtaataatac gagatgtttt 640
```

&lt;210&gt; 261

&lt;211&gt; 588

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:455450.1:2000SEP08

&lt;400&gt; 261

```
ggcggcgcca tcttcgccc ctccggcagc agctccagct ccagcagctc gtcacttgaa 60
gacgacggca tgggtggggc ccggaagaag ggcctgaagg agaagatcaa ggagaaaatg 120
ccaggaggcc acaggggaag ccagggccag gcgacggcca ccggtgcgta cggcgggaca 180
gggtacgtgg ctgggccgac gaccggaggc ccccacgaga agaaggggtg ggtggagaag 240
atcaaggaga agatcccagg cggccacaag gactacgacc agcatcagca caccaccgag 300
gcaaccgggt gcggcggtgg ctatggcgga accaccgaca ccacgtatgg gacgacgacc 360
actgaaggca ctacagagaa gaagggcttc atggacaaaa tcaaggagaa actccccggt 420
cagcactaag cactaagcag ctgcctcttt cctcttcggt ctagctgcct ctttcctctt 480
cgggttggcg actgcctgta agtacctgta tacagaataa ataaagcaat aaagaataag 540
cggcagctac gtatgtgtgt gtattcgtgt gtatgcattg taaggcag 588
```

&lt;210&gt; 262

&lt;211&gt; 383

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1073699.1:2000SEP08

&lt;400&gt; 262

```
gcttgccctg aggaactaag ctacagtctc tcgcggtcgc tggcctggat atggaccctg 60
agacctgccc ctgtcctact ggtggttcct gcacctgctc ggacaaatgc aaatgcaagg 120
gctgcaaatg caggaactgc aagaagagct gctgctcctg ttgccctgca ggatgtgaga 180
agtgtgccaa ggactgtgtt tgcaaaggcg aagagggggc caaggccgag aaatgcagct 240
gctgccagtg aggactccca cacagcctat gtgaatagtg ctgcgtgtcc ctggtggggc 300
gtggctgttg cccccctccc tggcttcctg ctccgggggt gtgaataaat cccatgcaca 360
gcatgaaaaa aaaaaaaaaa ggg 383
```

&lt;210&gt; 263

&lt;211&gt; 643

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: LI:1013729.1:2000SEP08

<400> 263

```
gtgctctccc gctcgccatg gccagcctta aggacctgga aggggaagtgg cgtctggtgg 60
aaagccacgg gtttgaggac tacatgaagg aactaggagt agggctggct cttaggaaga 120
tggttgccat ggccaaacca gactgcatca ttaccctcga cggcaacaac ctcaccgtca 180
aaactgagag cacggtgaag acgaccgtgt tttcttgac cttgggagag aagtttgatg 240
aaaccacagc tgatggcagg aaaactgaga cggctctgcac cttcacagac ggtgccctgg 300
tccagcacca gaagtgggaa gggaaagaaa gcacgataac gagaaaactg aaggacggga 360
agatgggtgg ggagtgcgtc atgaacaatt gccatcttg tacctcgggt ctatgacga 420
aggtacaatt gaggaactgg taactcgtca tcctggacag cagtcagctg gctgagggaa 480
taagctcaat tcaatgagca ggtcgtacag aaccacactg cttcacttct ttggttttat 540
ttttcatgac ttttcatcat agacacttta cccgaaacca tggtcagacg cgttggtttt 600
accaggaaca ttcctttggt ttagtacata aatgcgtttg tgc 643
```

<210> 264

<211> 1265

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:2050322.2:2000SEP08

<400> 264

```
gcggcgcggt gctgtacgag ggcggggccc aaaagcgcgt ggagcggcag aatgaactga 60
agcgcagggt tgagcagatg aagcaggacc ggctgaggcg gttaccaggg tgtgaacttg 120
tatgtgaaga atctggacga ctccattgat gacgacaaac tgaggaaaga gttctctccc 180
tatggagtaa ttaccagtgc gaaggtgatg acagagggtg gccacagcaa ggggtttggc 240
tttgtgtggt tttctctccc agaagaggcg acaaaggcca tgacagagat gaacggggcg 300
atcgtgggca ccaagccact ctacgtggca ctggccagc gcaaaagagg agcggaaggc 360
catcttgagc aaccagtaca tgcagcgctt ctccaccatg cggacccttg agcaaagccc 420
ctcctgggct cctttcagca gccctccagc tacttctgct ctgccatgcc ccagctgcct 480
accctccagg gtgcctcaat ggtccggcca accagggtgt ggctcggcg ccccccggcc 540
cacattcagc agtgtcaggc aggcctccac ccagggtgcca cgcacgggtg ctcataccca 600
gagagttagc aacattggta ctcagaccac aggaccagt ggggtaggat gctgtacacc 660
aggccggcgg ctctctgccc tgcaaatggt cctcagcagc acatacgcac ctatcgggta 720
aggtccagga agccggctgt gcacatccca gggacaggta gccctgatc cgcgtccatg 780
ctgagctatc ggcgcaccct gcattgagcaa aagcagatga ttgaagggtg tccacacagg 840
ggagcgtctc ctacccccct tatcccatga tggtcacacc ccaagggtggc tggcaagatc 900
cacgggcctt cgggttggga gattgacaag ctcagagctg ttgctacatg ctggagtctc 960
ccagaatccc tccatttcca agatagacga ggcagtggca gtgctagcag gcacaccaag 1020
gctatggagc agccgaaggc gtacatgcca ctgaaaacca gacaaaggaa atcctcgctt 1080
accatggctt gccaaaagga cagtgtttct ggctctcagc cctaaggccc tggaaaactc 1140
taacttattt cccaattagt ctgtatctat acttgggctc tgtatgtgaa tgacggtttg 1200
tcacccaacc agcctattac cttttgcttt gtgtattcaa agtgctgcaa aatctgcaaa 1260
aaaaa 1265
```

<210> 265

<211> 605

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:891327.1:2000SEP08

<400> 265

```
gaaagaagtg tatcctgcta cacctgcatg caccccaagc aaccgtctca cagctaaagg 60
agcagaggag actcttgga cttcagaaaa gaaaaaaacc atctgaagaa gagactggaa 120
ccaaaaggag taagatgtcc aaagagcaga ctcggccttc ctgctctgca ggagccagca 180
cgtccacagc catgggcccgt tccccacctc cccagacctc atcatcagct ccacccaaca 240
cttctctaac tgagagccta aaaccattgg ccaaccgtca cgcaactgcc agtaaaaaata 300
```

```

ttttccgaga agacccaata atcgcgatgg tactaaatgc aacaaaagta tttaaataatg 360
aatcctcaga aaatgagcaa agaagaatgt ttcattgctac agtgggctacg cagacacagt 420
tcttttcatgt gaagggtttta aacatcaact tgaagaggaa attcattaaa aagagaatca 480
tcattatatac aaattattcc aaacgtaata gtctctatag aggtgaatga agcctcttct 540
gtatctgaag ctgggtcctga ccaaacgttt gaggttccaa aggacatcat cagaagagca 600
aagaa 605

```

&lt;210&gt; 266

&lt;211&gt; 992

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2053076.1:2000SEP08

&lt;400&gt; 266

```

tgggaggaag gcagctgtaa agctacagat cgtaggcaag cctaacagag taagaaat 60
gaagctcact gtagtatttt agtttttagta gtagatctta gtcctactga tcgccatgtg 120
actcactttc tgcacttcct tctcttttagc atgttttctc ctcatctgga cttctgagga 180
cagctgggaa tgggtgaggaa gaaattctgc ccactgttaa cttgcagggt tatacctaca 240
cctgtgacgt ggggaagagg gagaacgtct acctgacggc tgaaagagtc cgcaaggagg 300
ttggcgaagt ctgagtcctg gtcaataatg ctgggtgtggt ctctggggca tcaactttct 360
ggaatgtcct gatgagctca ttgagagaac catgatggtc aattgccatg cacacttctg 420
gaccactaag gcttttcttc ctacgatgct cggagattaa tcatggtcac attgtgacag 480
ttgcaagttc cttgggattg ttcagtactg ccggagttga ggattactgt gccagtaaat 540
ttggagttgt gggttttcat gatatccctg agccatgaac taaaggctgc tgacaaagga 600
tggaattaaa acaaccttgg tttgccctta tctttagtag actggcatgt tcagaggctg 660
ccgaatcagg aaagaaattg agccttttct gccaccttct gaagcctgat tactgtgtga 720
agcaggccat gaaggccatc ctactgacc agcccatgat ctgcaactccc cgcctcatgt 780
acatcgtgac cttcatgaag agcatctac catttgaagc agttgtgtgc atgtatcgg 840
tcctaggagc ggacaagtgt atgtaccctt ttattgtctca tatagaaagc aagccacaaa 900
caataatgaa gcaataaat ggaatctaag aatctttttg tatggaatat tacttctatc 960
agaagatgat caagatgttt cagtccagtg ca 992

```

&lt;210&gt; 267

&lt;211&gt; 928

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:220085.1:2000SEP08

&lt;400&gt; 267

```

gctcccctag agccagtggg acccttcggg accctgggaa aaacactcag tctgaccaac 60
catgcctgca gcataaccag tgtgcatact gtaaagaaac agcgcattgg aaagacaagt 120
gcccctagtt aaaagagaaa caaaatggct ctgagccaga ggtttcaggc aaggatgaag 180
gggccttggt taatctagca gaagggttac tggactgagg gggaccgggc tcagggtgcc 240
ccaatgagcc catggtcagg atgacagtca ggagcaagga cattgagttt cctgtcgata 300
cttatcctga acattcagta gtaaccacc cagtcacccc cttatccaaa aagactattg 360
atataatcaa agccacagga gtcttgaaaa aacaagcttt ctacttgccc cggacttgta 420
ctgtagggag acatgaagtg attcaccagt tcttgtagat gcctgactgc cccttgccct 480
tgctgggaag ggacttgctt agcaagctga gggccacat ctcttttaca aagcacagct 540
ctttaccgct aaagttaacct ggaatgggag tcattatggc ccttacagtc ccctgggagg 600
aagaatggac ttttcttaac agccaggcca agagatagga ccagcttttg ctaaacagtg 660
gctaagggtg tgggcagaag ataactctct ggggctagca gtaaacgaag cccccgtaac 720
cagagaagct cttgaaggta tccagatcca tctccagtgc ctgaaggcct ttggaattat 780
agccccctgt cagtctgcct ggaacacagc cctgctgcct gttcccaatc caggaaccaa 840
ggtttacaga ccagtatagg acttgcgctt ggttaaccaa gctacagtga ctttgacccc 900
aatggtacct aacccgtaca tgttctta 928

```

&lt;210&gt; 268

<211> 1323  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:406709.1:2000SEP08

<400> 268  
ggcagctcct ggaagctcca ccaggaccca cgtccaccag gcatctcgaa cccttggccc 60  
ccatggaccc cctgagggcg tccccctctc cctgcctgtc ctgcagagcc tccagcccca 120  
gcacccacc ctgcgagatg cttggctcctg tgggcattga ggctgtgctg gaccagctga 180  
agatcaaggc tatgaagatg gggtttgagt tcaacatcat ggtgggtggg caaagcgggc 240  
tgggcaagtc cacgatgggt aacacgctgt tcaagtccaa agtgtggaag tcaaaccac 300  
cgggcttggg ggtgccaca cccagacgc tgcagctgca ttcactgacc catgtcatag 360  
aggagaaggg tgtgaagctg aagctgacgg tgacggacac gcccggttc ggggaccaga 420  
tcaacaatga caactgctgg gaccccatcc tgggctacat caacgagcaa tacgagcagt 480  
acctgcagga ggagatcctc atcaccgcc agcgccacat cccagacacc cgggtgacct 540  
gctgcgtgta ctttgtacca cccactgggc actgcctgcg gccctggac attgagttcc 600  
tgcagcgggt gtgtccggac tgtgaatgtg gtgcccgtga ttgccagggc cagacagcct 660  
gacctggag gagcgagagg ccttcaggcg caggatccag cagaacctga ggaccactg 720  
catcgacgtc tccccccaga tgtgctttga cgaggacatc aatgacaaaa tcctcaacag 780  
caagttacgg gaccgaatcc cttttgccgt ggtaggggct gaccaagagc acctggtgaa 840  
cgggaggtgt gtccctgggc ggaagaccaa gtggggcatc attgaagtgg agaacatggc 900  
gcactgtgaa tttcctctcc tgagagacct gcttatccgc tcccacctcc aagacctgaa 960  
ggacataacc cacaacatcc actatgagaa ctaccgcgtc atcagactca atgaaagcca 1020  
cctgctgccc cgcgggcccc gctgggtgaa cctggccccg gcctccccag gacagctgac 1080  
caccccccg accctcaagg tctgcagggg ggcccatgac gattctgatg atgagttctg 1140  
accaccggcg gatcccgagg ctgctgggct tcctgagtc ccagcggctc tcaacacaca 1200  
cctatgtacc agagcatcta ttaaagtga gccttgcttt ttatgaaaag ctgtgctttg 1260  
aaaacaaaag gcattttgta aatgacttct ttgagctatc cacaaataaa aaggctgggt 1320  
gtc 1323

<210> 269  
<211> 745  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LG:347863.9:2000SEP08

<400> 269  
gggcatgtat ggatgtgtgg atattctatg tctgtattta gactgcaggt atatggatgt 60  
tcatgtgctg tgtgtgtgtg tgcgcacaca cactctgcct ctttgtgtgt gtgcatgtgt 120  
ataccttgtg tgcccatgta tcgggggtgt gtgtaccctg cttgtctatg tatgggggag 180  
cacatgtagc ctgtgagttc ttattctgct gggcaggtgt gtggaggtgg tgggggaggc 240  
tgagcatgaa aggccacttg tgccagttgg aggaggggtg gcgctgctgc ctgctgcctc 300  
aggacttggt gggcctgtgg gaggccgggc tccaccacc agccttcctt gttctcggtc 360  
cccagtttaa gaggcgagtt caggagtcca cgcaggtgct acgggagctg gagacctccc 420  
tgaggaccaa ccacattggg tgggtgcagg agttcctcaa tgaagagaac cgtggcctgg 480  
atgtgctgct cgagtacctg gcctttgccc agtgcctgt cactgaagcc ccctgctccc 540  
agccctcatg ccgctcctca gagctttgat ccccgctctc ctgcatctca cccactcccc 600  
tggccagttt caagccaggc agcccagacc taccctggaa ccctccactt ggccttgagc 660  
gatgctcctt ccagaaggcc tgccccgcac agggaggggt ggcctctctt ccaccactat 720  
gttcagcaca gtgccaaagaa cacag 745

<210> 270  
<211> 582  
<212> DNA  
<213> Homo sapiens

<220>

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1073027.1:2000SEP08

&lt;400&gt; 270

```
atagaaatgg catccaaaag agctctgggtc atcctagcca aaggagcaga ggagatggag 60
acagtgattc ctgtggacat catgcccga gctgggatta aagtcaccgt tgcaggcttg 120
gctgggaagg accccgtgca gtgtagccgt gatgtagtga tttgtccgga taccagtctg 180
gaagaagcaa aaacacaggg accatacgat gtggttgttc ttccaggagg aaatctgggt 240
gcacagaact tatctgagtc ggctttgggtg aaggagatcc tcaaggagca ggagaacagg 300
aagggcctca tagctgccat ctgtgcccgt cctacggccc tgcaggctca cgaagtaggc 360
tttggatgca aggttacatc gcaccattg gctaaggaca aaatgatgaa cggcagtcac 420
tacagctact cagagagccg tgtggagaag gacggcctca tcctcaccag ccgtgggcct 480
gggaccagct tgcagtttgc gctggccatt gtggaccctg gacccccagg ctgagcaggc 540
attggaagcc cactagtgtg tccacagcac agtgaacgtc ag 582
```

&lt;210&gt; 271

&lt;211&gt; 1717

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:347635.1:2000SEP08

&lt;400&gt; 271

```
gccagaggg aagcatgtct cctgtgtcat ctggatgatt gcatgcatca gcaatccttg 60
ccacaagggg gcaactgtgt acaccaaccc cctaaatggg caatatatatt gcacctgccc 120
acaaggctac aaaggggctg actgcacaga agatgtggat gaatgtgcca tggccaatag 180
caatccttgt gagcatgcag gaaaatgtgt gaacacggat ggcgccttcc actgtgagtg 240
tctgaagggt tatgcaggac ctgctgttga gatggacatc aatgagtgcc attcagacct 300
ctgcagaat gatgtacct gtctggataa gattggaggc ttcacatgtc tgtgcatgcc 360
agggtttcaaa ggtgtgcatt gtgaattaga aataaatgaa tgcagagca acccttgtgt 420
gaacaatggg cagtgtgttg ataaagtcaa tcgtttccag tgcctgtgtc ctccctgggtt 480
cactgggcca gtttgccaga ttgatattga tgactgttcc agtactccgt gtctgaatgg 540
ggcaaagtgt atcgatcacc cgaatggcta tgaatgccag tgtgccacag gtttactggt 600
tgtgttgtgt gaggagaaca ttgacaactg tgaccccgat ccttgccacc atggtcagtg 660
tcaggatggt attgattcct acacctgcct ctgcaatccc ggggtacatgg gcgccaatct 720
gcagtgacca gattgatgaa tgttacagca gcccttgccct gaacgatggg cgctgcattg 780
acctgggtcaa tggctaccag atgcaactgc cagccaggca cgtcaggggg taattgtgaa 840
attaattttg atgactgtgc aagtaaccct tgtatccatg gaatctgtat ggatggcatt 900
aatcgctaca gttgtgtctg ctcaccagga ttccacaggc cagagatgta acattgacat 960
tgatgagtgt gcctccaatc cctgtcgcaa ggggtgcaaca tgtatcaacg gtgtgaatgg 1020
tttccgctgt atatgcccgg agggacccca tcaccccagc tgctactcac aggtgaacga 1080
atgcctgagc aatccctgca tccatggaaa ctgtactgga ggtctcagtg gatataagtg 1140
tctctgtgat gcaggctggg ttggcatcaa ctgtgaagtg gacaaaaatg aatgcctttc 1200
gaatccatgc cagaatggag gaacttgtga caatctgggt aatggataca ggtgtacttg 1260
caagaagggc tttaaaggct ataacatgcc aggtgaatat tgatgaatgt gcctcaaate 1320
catgcctgaa ccaaggaacc tgctttgatg acataagtgg ctacacttgc cactgtgtgc 1380
tgccatacac aggtgggtcc tggaggtacc agcagggacc agaggggata cttgaccttg 1440
ctttgtcagt gcctgagtct cagtgccagg gtctgtgggg ctgtaatgct cctccgaggg 1500
ctcactcctc agttgtcctt cctgccttta ctgactggat cttgagcttt gaggacacat 1560
tgtggcctcc taagatgcaa gtggtggaca gtcatttgcc tacattgtac tttagtattt 1620
gaaagaaatt ttctcttaat tgtatgtggc tggaaccagg cacatagatg atggttgtct 1680
atcatttgaa tgtttccagg acattagcta tttgtct 1717
```

&lt;210&gt; 272

&lt;211&gt; 2218

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:013685.1:2000SEP08

&lt;400&gt; 272

```
caggcgtgag ccaccacgcc cagcctcctt tttcacttcc taccttttct cccacacctc 60
ttgggttgta ccctccccgt gagctcagcc tgtatttcta tctgctggcc aacttaccag 120
ccagtccttt catctacaat gccctttact caccatagaa gctcacttca aaagggtcgcc 180
tcctccagaa tacatctcaa gttctcctct ctgggtccctt gaccttttgc aggtaccttt 240
tttctggac caagtaacaa taataatgta aggggttttt ttccccatct agagcctgcc 300
tgggtgggctc atgggtocatt tgtggaaaaa tgcatttccc tctcctccgc ccacctcctt 360
catgatccat gctttaagag ctcccccaag gcatgcagag cagagctgag aagcagggtta 420
cgacaagcct gcagaccttt caaagaagtg ggggtggctt ggagtaaaaa gaagatctgg 480
agataggagt tcctaaaaag cagacaagag attaccgcag gggttcccca cagctccagc 540
acacagacag ctgcaaaaaa aatgcagtgg aattactgag ggacattatg ggcttgatgg 600
aggacagggg atgcttggtgc caagccacca aggatcatgc tgtggcagtt gtgaccagat 660
agactatact ttgccaagg tgaaggatt tcctactaa atcctttact ttccactttg 720
tttaaaatga gtttaacacc aagaaaatag ttgtctctat gtccttattt ctaatcttcc 780
tgtcaacca tgaactgcac taaacacact cagaccatc taacctaaaa ttgagaaagg 840
gaggtttctca ctctggaga gagaattact gaaagggcgg ccctcataca gttcagggtc 900
atacccccac gagctgggtt tatatttgaa agcagacact gggactctga cctttgcaga 960
gataaaaggg atacttttct aaccagaga ggccaaagga gcacaggacg ccagggccga 1020
gggcagtga gcttgacagc ccaggctgtg caacccccac agcagggtctc ctcacgggat 1080
agatagacta ttggcccaa ttctcagcc ctacatgcac ccacctttg ccatgggttc 1140
gttggtgggtg gaggttactt tcccagcccc tggctttggg cttgcccacg tgtcttgctc 1200
tggccatgga atgaagcaca aaggagagcc tgctagtcc taacccacgc ctgaagacgc 1260
cttaggcggt tctgtcggac ccctgtcact tccaccttca gccagaaggc cttctctggt 1320
gcagccgttg cttcagcctc ggtccaataa ggatcggagc ccctgtccg atgccgtcag 1380
gctgtcagg gagctcatca tgagcgtcag ctggcccgcc ccagcgttca agccaatgt 1440
cacagtgcct gcaaggaaact ttctttccaa ctgcctgttt tgggagtggg ttgattagga 1500
cgctcgttgcc tgtggcggta gctcaccaat ccaatgcag catcccgctc ctttatttag 1560
gctatatggt ccatgtggct tgcgacaggg acctgataca tacatgtgcg ttgaaataag 1620
gagcttattg agtctatcat gtcataagtc cgttgagaa tgtccaggtc ctagtgtagg 1680
gtctctgtcc ggggtctgta gccctcatcc gtcgtggaac cttcccatgg ttgcgagtc 1740
gcgggtctcg gagccttcag ccctcgtgag ccgatacag tccatagggc ggcgcccagc 1800
cctcttccct acgccacct cttaaagcctc ggctccaacc gattccactt ctgcttcagg 1860
ctcaggattt tcaactctct cgaatggggg tggccctccc ctgatcttct gaatctgcaa 1920
cagcatctcc ctccccccag gacatcagag ccagagctgg gcgagaggcc ctgacctctg 1980
gggtaggggt gcagcgtccc cttgaagatg cggctccgccc tcccatcccc aggcgcccgg 2040
cctctcacac cctcagcacc ctgctcacct ccagctgaag atgccagggc acctctgctt 2100
cctgctgct ctctgcagta ccgcagagtg agcataaaag ggcttaatac aggccttcgcc 2160
gggcgcaggg ctcccgccctg taattccagt actttgggag accaaggcgg gaggggtca 2218
```

&lt;210&gt; 273

&lt;211&gt; 1330

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:406709.1:2000SEP08

&lt;400&gt; 273

```
ggcagctcct ggaagctcca ccaggaccca cgtccaccag gcatctcgaa cccttggccc 60
ccatggaccc cctgaggcgc tccccctctc cctgcctgtc ctgcgagccc tccagcccca 120
gcaccccacc ctgcgagatg cttggtcctg tgggcattga ggctgtgctg gaccagctga 180
agatcaaggc tatgaagatg gggtttgagt tcaacatcat ggtggtgggg caaagcgggc 240
tgggcaagtc cacgatgggtg aacacgctgt tcaagtccaa agtgtggaag tcaaacccac 300
cgggcttggg ggtgcccaca cccagacgc tgcaagctgc attcactgac ccatgtcata 360
gaggagaagg gtgtgaagct gaagctgac gtgacgggac acgcccggct tcggggacca 420
gatcaacaat gacaactgct gggaccccat cctgggtac atcaacgagc aatacagca 480
gtacctgcag gaggagatcc tcatcacccg ccagcgccac atcccagaca cccgggtgca 540
ctgctgcgtg tactttgtac caccactgg gcactgcctg cggcccctgg acattgagtt 600
cctgcagcgg ctgtgtccgg actgtgaatg ttgtgcccgt gattgccagg gccagacagc 660
ctgatccatg gaggagcag aggccttcag gcgcaggatc cagcagaacc tgaggacca 720
ctgcatcgac gtcctacccc agatgtgctt tgacgaggac atcaatgaca aaatcctcaa 780
cagcaagtta cgggaccgaa tcccttttgc cgtggtaggg gctgaccaag agcacctggt 840
```

```

gaacgggagg tgtgtcactg ggccggaaga ccaagtgggg catcattgaa gtggagaaca 900
tggcgcaactg tgaatttcct ctcttgagag acctgcttat ccgtcccccac ctccaagacc 960
tgaaggacat aacccacaac atccactatg agaactaccg cgtcatcaga ctcaatgaaa 1020
gccacactgc tgccccgcgg gcccggtcgg gtgaacctgg ccccgccctc cccaggacag 1080
catgaccacc ccccgacct tcaaggatcat gcaggggggc ccatgacgat tctgatgatg 1140
agttctgacc accggcggtat cccggggctg ctgggcttcc tgagtcccca gcggctctca 1200
acacacacct atgtaccaga gcatctatta aatgtgagcc ttgcttttta tgaaaagctg 1260
tgctttgaaa acaaaaggca ttttgtaaat gacttctttg agctatccac aaataaaaaa 1320
gctgggtgtc 1330

```

&lt;210&gt; 274

&lt;211&gt; 747

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2052938.1:2000SEP08

&lt;400&gt; 274

```

cattgcattg agaggatatt caaactgaag gggagaagca cggattctga aggatgaaaa 60
aggaacaatg tcgttttctg atgtgacggc tgagacatga gatcttcagc ctccaggctc 120
tccagttttt cgtcgagaga ttcactatgg aatcgatata caggaatgat ttccagagta 180
tttgagaaaa tgagaggtgc agatggtttt cactgaggca tggtttcaaa tttaaagagg 240
attatggccc agaaatggat gccggaccag atctctgtct cggagttcat cgccgagacc 300
accgaggact acaactcgcc caccacgtcc agcttcacca cgcggctgca caactgcagg 360
aacacogtca cgctgctgga ggaggctcta gaacaagata gaacagccct tcagaaagt 420
aagaagtctg taaaagcaat atataattct ggtcaagatc atgtacaaaa tgaagaaaac 480
tatgcacaag ttcttgataa gtttgggagt aattttttaa gtcgagacaa ccccgacact 540
tggcaccgag tttgtcaagt tttctactct tacaaggaa ctgtccacac tgctgaaaaa 600
tctgtctcag ggtttgagcc cacaatgtga tcttcacctt ggattctttg ttaaaaggag 660
acctaaaggg agtcaaagga gatgctcaag aagccatttg acaaagcctg gaaagattat 720
gagacaaagt tacaaaactg ggaaaaa 747

```

&lt;210&gt; 275

&lt;211&gt; 498

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:213208.1:2000SEP08

&lt;400&gt; 275

```

cgagcgagat ggggcggagg gccggtcttg cagagtagct gcggtgagtg ggcgtgtgag 60
ccgagcggtc tggcccaagg gctggggggc gcccgagggt ctccgggagc aggcgcagg 120
gcgcggagag atcctgggat cgcgctccgc cgctgctacc cggcatgtcg gcggaggcct 180
cgggcccgcc tgccgcgcgg gcccgctccc tggaaagccc caagccctcg ggtctcgagc 240
ctggccccgc cgctacggt ctcaagccgc tgaccccgaa cagcaaatac gtgaagtga 300
acgtggggcg ctggttgac tacaccacgc tgcgcaccct cacgggacag gacaccatgc 360
tcaaagccat gttcagcgcc cgcgtggagg tgctgaccga tgccggagggt tgggtgtgta 420
ttgaccggag cggccgtcac tttggtacaa tcctcaatta cctgcgggat ggggtctgtg 480
cactgccgga gagtacga 498

```

&lt;210&gt; 276

&lt;211&gt; 133

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:405741.3.orf3:2000SEP08



&lt;400&gt; 276

```

Thr Asp Cys Glu Gly Glu Asp Val Val Asp Met Leu Arg Glu Ala
 1          5          10          15
Ile Lys Arg Arg Asn Val Gly Cys Gly Val Glu Ala His Ala Cys
          20          25          30
Ser Cys Cys Leu Pro Gln Ala Pro Leu Leu Cys His Pro Leu Pro
          35          40          45
Ser Ser Tyr Asp Ser Gly Pro Ser Pro Gly Gln Glu Asp Ser Ala
          50          55          60
Ala Ala Trp Ala Trp Thr Ala Arg Val Arg Ile Thr Val Gly Val
          65          70          75
Asp Gly Thr Leu Tyr Lys Leu His Pro His Phe Ser Arg Ile Leu
          80          85          90
Gln Glu Thr Val Lys Glu Leu Ala Pro Arg Cys Asp Val Thr Phe
          95          100          105
Met Leu Ser Glu Asp Gly Ser Gly Lys Gly Ala Ala Leu Ile Thr
          110          115          120
Ala Val Ala Lys Arg Leu Gln Gln Ala Gln Lys Glu Asn
          125          130

```

&lt;210&gt; 277

&lt;211&gt; 160

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:337194.1.orf1:2000SEP08

&lt;400&gt; 277

```

Phe Val Ala Tyr Leu Lys Leu Pro Phe Phe Ser Gln Val Cys Leu
 1          5          10          15
Phe Ala Ser Ser Glu Met Phe Phe Thr Ile Ser Arg Lys Asn Met
          20          25          30
Ser Gln Lys Leu Ser Leu Leu Leu Leu Val Phe Gly Leu Ile Trp
          35          40          45
Gly Leu Met Leu Leu His Tyr Thr Phe Gln Gln Pro Arg His Gln
          50          55          60
Ser Ser Val Lys Leu Arg Glu Gln Ile Leu Asp Leu Ser Lys Arg
          65          70          75
Tyr Val Lys Ala Leu Ala Glu Glu Asn Lys Asn Thr Val Asp Val
          80          85          90
Glu Asn Gly Ala Ser Met Ala Gly Tyr Ala Asp Leu Lys Arg Thr
          95          100          105
Ile Ala Val Leu Leu Asp Asp Ile Leu Gln Arg Leu Val Lys Leu
          110          115          120
Glu Asn Lys Val Asp Tyr Ile Val Val Asn Gly Ser Ala Ala Asn
          125          130          135
Thr Thr Asn Gly Thr Ser Gly Asn Leu Val Pro Val Thr Thr Asn
          140          145          150
Lys Arg Thr Asn Val Ser Gly Ser Ile Arg
          155          160

```

&lt;210&gt; 278

&lt;211&gt; 125

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:017108.4.orf2:2000SEP08

&lt;400&gt; 278

```

Ala Met Trp Leu Cys Lys His Leu Asn Ser Ser Leu Leu Thr Leu
  1          5          10          15
Glu Asn Leu Ile Leu Asn Glu Phe Ser Tyr Thr Ala Thr Glu Ala
          20          25          30
Arg Arg Leu Tyr Leu Gln Arg Lys Thr Val Pro Ser Ala Leu Leu
          35          40          45
Val Gln Leu Ile Gln Glu Arg Leu Ala Glu Glu Asp Cys Ile Lys
          50          55          60
Gln Gly Trp Ile Leu Asp Gly Ile Pro Glu Thr Arg Glu Gln Ala
          65          70          75
Leu Arg Ile Gln Thr Leu Gly Ile Thr Pro Arg His Val Ile Val
          80          85          90
Leu Ser Ala Pro Asp Thr Val Leu Ile Glu Arg Asn Leu Gly Lys
          95          100          105
Arg Ile Asp Pro Gln Thr Gly Glu Ile Tyr His Thr Thr Phe Asp
          110          115          120
Trp Pro Pro Glu Ser
          125

```

&lt;210&gt; 279

&lt;211&gt; 373

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:372569.5.orf1:2000SEP08

&lt;400&gt; 279

```

Gly His Cys Pro Cys Ser Met Thr Arg Ser Thr Pro Glu Pro Ala
  1          5          10          15
Ala Ser Leu Arg Thr Pro Trp Val Leu Pro Pro Ala Ser Arg Asn
          20          25          30
Gln Arg Ile Leu Tyr Thr Val Leu Glu Cys Gln Pro Leu Phe Asp
          35          40          45
Ser Ser Asp Met Thr Ile Ala Glu Trp Val Arg Val Ala Gln Thr
          50          55          60
Ile Lys Arg His Tyr Glu Gln Tyr His Gly Phe Val Val Ile His
          65          70          75
Gly Thr Asp Thr Met Ala Leu Ala Ala Ser Met Leu Ser Phe Met
          80          85          90
Leu Glu Asn Leu Gln Lys Thr Val Ile Leu Thr Gly Ala Gln Val
          95          100          105
Pro Ile His Ala Leu Trp Ser Asp Gly Arg Glu Asn Leu Leu Gly
          110          115          120
Ala Leu Leu Met Ala Gly Gln Tyr Val Ile Pro Glu Val Cys Leu
          125          130          135
Phe Phe Gln Asn Gln Leu Phe Arg Gly Asn Arg Ala Thr Lys Val
          140          145          150
Asp Ala Arg Arg Phe Ala Ala Phe Cys Ser Pro Asn Leu Leu Pro
          155          160          165
Leu Ala Thr Val Gly Ala Asp Ile Thr Ile Asn Arg Glu Leu Val
          170          175          180
Arg Lys Val Asp Gly Lys Ala Gly Leu Val Val His Ser Ser Met
          185          190          195
Glu Gln Asp Val Gly Leu Leu Arg Leu Tyr Pro Gly Ile Pro Ala
          200          205          210
Ala Leu Val Arg Ala Phe Leu Gln Pro Pro Leu Lys Gly Val Val
          215          220          225
Met Glu Thr Phe Gly Ser Gly Asn Gly Pro Thr Lys Pro Asp Leu
          230          235          240
Leu Gln Glu Leu Arg Val Ala Thr Glu Arg Gly Leu Val Ile Val
          245          250          255

```

Asn	Cys	Thr	His	Cys	Leu	Gln	Gly	Ala	Val	Thr	Thr	Asp	Tyr	Ala	
				260					265					270	
Ala	Gly	Met	Ala	Met	Ala	Gly	Ala	Gly	Val	Ile	Ser	Gly	Phe	Asp	
				275					280					285	
Met	Thr	Ser	Glu	Ala	Ala	Leu	Ala	Lys	Leu	Ser	Tyr	Val	Leu	Gly	
				290					295					300	
Gln	Pro	Gly	Leu	Ser	Leu	Asp	Val	Arg	Lys	Glu	Leu	Leu	Thr	Lys	
				305					310					315	
Asp	Leu	Arg	Gly	Glu	Met	Thr	Pro	Pro	Ser	Val	Glu	Glu	Arg	Arg	
				320					325					330	
Pro	Ser	Leu	Gln	Gly	Asn	Thr	Leu	Gly	Gly	Gly	Val	Ser	Trp	Leu	
				335					340					345	
Leu	Ser	Leu	Ser	Gly	Ser	Gln	Glu	Ala	Asp	Ala	Leu	Arg	Asn	Ala	
				350					355					360	
Leu	Val	Pro	Ser	Leu	Ala	Trp	Cys	Cys	Cys	Pro	Arg	Arg			
				365					370						

&lt;210&gt; 280

&lt;211&gt; 227

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:968765.1.orf2:2000SEP08

&lt;400&gt; 280

Arg	Ile	Ala	Tyr	Gly	Gly	Arg	Ser	Gly	Gly	Tyr	Ser	Ser	Gly	Lys	
1				5					10					15	
Gln	Gly	Arg	Gly	Asp	Ser	Ser	Ser	Lys	Lys	Asp	Val	Val	Glu	Leu	
				20					25					30	
Thr	Asp	Arg	His	Leu	Met	Ile	Arg	Met	Ser	Leu	Thr	Val	Lys	Thr	
				35					40					45	
Phe	Gly	Trp	Leu	Ser	Phe	Ile	Ala	Pro	Trp	Cys	Gly	His	Cys	Lys	
				50					55					60	
Asn	Leu	Glu	Pro	Glu	Trp	Ala	Ala	Ala	Ala	Thr	Glu	Val	Lys	Glu	
				65					70					75	
Gln	Thr	Lys	Gly	Lys	Val	Lys	Leu	Ala	Ala	Val	Asp	Ala	Thr	Val	
				80					85					90	
Asn	Gln	Val	Leu	Ala	Ser	Arg	Tyr	Gly	Ile	Lys	Gly	Phe	Pro	Thr	
				95					100					105	
Ile	Lys	Ile	Phe	Gln	Lys	Gly	Glu	Ser	Pro	Val	Asp	Tyr	Asp	Gly	
				110					115					120	
Gly	Arg	Thr	Arg	Ser	Asp	Ile	Val	Ser	Arg	Ala	Leu	Asp	Leu	Phe	
				125					130					135	
Ser	Asp	Asn	Ala	Pro	Pro	Glu	Leu	Leu	Glu	Ile	Ile	Asn	Glu		
				140					145					150	
Asp	Ile	Ala	Lys	Lys	Thr	Cys	Glu	Glu	His	Gln	Leu	Cys	Val	Val	
				155					160					165	
Ala	Val	Leu	Pro	His	Ile	Leu	Asp	Thr	Gly	Ala	Thr	Gly	Arg	Asn	
				170					175					180	
Ser	Tyr	Leu	Glu	Val	Leu	Leu	Lys	Leu	Ala	Asp	Lys	Tyr	Lys	Lys	
				185					190					195	
Lys	Met	Trp	Gly	Trp	Leu	Trp	Thr	Glu	Ala	Gly	Ala	Gln	Tyr	Glu	
				200					205					210	
Leu	Glu	Asn	Ala	Leu	Gly	Ile	Gly	Gly	Phe	Gly	Tyr	Pro	Gly	Met	
				215					220					225	

Ala Ala

&lt;210&gt; 281

&lt;211&gt; 157

&lt;212&gt; PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:255999.16.orf3:2000SEP08

<220>

<221> unsure

<222> 138

<223> unknown or other

<400> 281

Thr	Leu	Ala	Gly	Leu	Gln	Asp	Arg	Asn	Val	Ala	Lys	Arg	Val	Leu
1				5					10					15
Cys	Ile	Val	Pro	Phe	Arg	Leu	Val	Ser	Gln	Val	Asn	Val	Asp	His
				20					25					30
Phe	Met	Ser	Asp	Val	Phe	Gly	Leu	Glu	Asp	Glu	Ala	Gly	Pro	Val
				35					40					45
Arg	Ile	Gly	Ala	Glu	Pro	His	Ala	Val	Asp	Ala	Asp	Gln	Pro	Leu
				50					55					60
Arg	Asp	Trp	Pro	Trp	Gly	Leu	Ala	Ser	Trp	Asp	Leu	Glu	Gly	Ser
				65					70					75
Val	Gly	Lys	Cys	Gly	Arg	Glu	Gly	Ala	Gly	Gly	Ser	Pro	Gly	Gly
				80					85					90
Gln	Ala	Cys	Val	Gly	Leu	Ser	Cys	Glu	His	Thr	Arg	Ser	Cys	Pro
				95					100					105
Arg	Gly	Gln	Arg	Leu	Gly	Asp	Ala	Asp	Lys	Arg	Leu	Gly	Trp	Arg
				110					115					120
Val	Leu	Gly	Ser	Arg	Thr	Ala	Pro	Gly	Gly	Val	Gly	Ala	Val	Arg
				125					130					135
Arg	Gln	Xaa	His	Pro	Pro	Leu	Leu	Cys	Gly	Pro	Gly	Arg	Gly	Ser
				140					145					150
Gly	Gln	Pro	Gly	Ser	Leu	Pro								
				155										

<210> 282

<211> 281

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:977820.9.orf2:2000SEP08

<400> 282

Ser	Arg	Val	Leu	Ala	Gly	Arg	Ile	Asn	Asp	Ala	Lys	Cys	Lys	Val
1				5					10					15
Val	Ile	Thr	Phe	Asn	Gln	Gly	Leu	Arg	Gly	Gly	Arg	Val	Val	Glu
				20					25					30
Leu	Lys	Lys	Ile	Val	Asp	Glu	Ala	Val	Lys	His	Cys	Pro	Thr	Val
				35					40					45
Gln	His	Val	Leu	Val	Ala	His	Arg	Thr	Asp	Asn	Lys	Val	His	Met
				50					55					60
Gly	Asp	Leu	Asp	Val	Pro	Leu	Glu	Gln	Glu	Met	Ala	Lys	Glu	Asp
				65					70					75
Pro	Val	Cys	Ala	Pro	Glu	Ser	Met	Gly	Ser	Glu	Asp	Met	Leu	Phe
				80					85					90
Met	Leu	Tyr	Thr	Ser	Gly	Ser	Thr	Gly	Met	Pro	Lys	Gly	Ile	Val
				95					100					105
His	Thr	Gln	Ala	Gly	Tyr	Leu	Leu	Tyr	Ala	Ala	Leu	Thr	His	Lys
				110					115					120
Leu	Val	Phe	Asp	His	Gln	Pro	Gly	Asp	Ile	Phe	Gly	Cys	Val	Ala
				125					130					135

Asp	Ile	Gly	Trp	Ile	Thr	Gly	His	Ser	Tyr	Val	Val	Tyr	Gly	Pro	
				140					145					150	
Leu	Cys	Asn	Gly	Ala	Thr	Ser	Val	Leu	Phe	Glu	Ser	Thr	Pro	Val	
				155					160					165	
Tyr	Pro	Asn	Ala	Gly	Arg	Tyr	Trp	Glu	Thr	Val	Glu	Arg	Leu	Lys	
				170					175					180	
Ile	Asn	Gln	Phe	Tyr	Gly	Ala	Pro	Thr	Ala	Val	Arg	Leu	Leu	Leu	
				185					190					195	
Lys	Tyr	Gly	Asp	Ala	Trp	Pro	Gly	Met	Ala	Arg	Thr	Ile	Tyr	Gly	
				200					205					210	
Asp	His	Gln	Arg	Phe	Val	Asp	Ala	Tyr	Phe	Lys	Ala	Tyr	Pro	Gly	
				215					220					225	
Tyr	Tyr	Phe	Thr	Gly	Asp	Gly	Ala	Tyr	Arg	Thr	Glu	Gly	Gly	Tyr	
				230					235					240	
Tyr	Gln	Ile	Thr	Gly	Arg	Met	Asp	Asp	Val	Ile	Asn	Ile	Ser	Gly	
				245					250					255	
His	Arg	Leu	Gly	Thr	Ala	Glu	Ile	Glu	Asp	Ala	Ile	Val	Ser	Thr	
				260					265					270	
Gly	Trp	Ala	Ser	Ser	Phe	Thr	Pro	Thr	Pro	Pro					
				275					280						

&lt;210&gt; 283

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1071608.1.orf3:2000SEP08

&lt;400&gt; 283

Ile	Leu	Ala	Thr	Leu	Ser	Ser	Pro	Gln	Pro	Val	Leu	Ile	Ala	Lys	
1				5					10					15	
Gly	Thr	Met	Thr	Ala	Ser	Ser	Val	Leu	Leu	His	Thr	Gly	Gln	Lys	
				20					25					30	
Met	Pro	Leu	Ile	Gly	Leu	Gly	Thr	Trp	Lys	Ser	Glu	Pro	Gly	Gln	
				35					40					45	
Val	Lys	Ala	Ala	Ile	Lys	Tyr	Ala	Leu	Ser	Val	Gly	Tyr	Arg	His	
				50					55					60	
Ile	Asp	Cys	Ala	Ser	Val	Tyr	Gly	Asn	Glu	Thr	Glu	Ile	Gly	Glu	
				65					70					75	
Ala	Leu	Lys	Glu	Ser	Val	Gly	Ser	Gly	Lys	Ala	Val	Pro	Arg	Glu	
				80					85					90	
Glu	Leu	Phe	Val	Thr	Ser	Lys	Leu	Trp	Asn	Thr	Lys	His	His	Pro	
				95					100					105	
Glu	Asp	Val	Glu	Pro	Ala	Leu	Arg	Lys	Thr						
				110					115						

&lt;210&gt; 284

&lt;211&gt; 214

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1074023.1.orf3:2000SEP08

&lt;400&gt; 284

Val	His	Thr	Ser	Val	Tyr	Ala	Ala	Ala	Met	Pro	Pro	Tyr	Thr	Ile	
1				5					10					15	
Val	Tyr	Phe	Pro	Val	Arg	Gly	Arg	Cys	Glu	Ala	Thr	Arg	Met	Leu	
				20					25					30	
Leu	Ala	Asp	Gln	Gly	Gln	Ser	Trp	Lys	Glu	Glu	Val	Val	Thr	Ile	

	35		40		45
Asp Val Trp Leu Gln Gly Ser Leu Lys Ser Thr Cys Leu Tyr Gly					
	50		55		60
Gln Leu Pro Lys Phe Glu Asp Gly Asp Leu Thr Leu Tyr Gln Ser					
	65		70		75
Asn Ala Ile Leu Arg His Leu Gly Arg Ser Leu Gly Leu Tyr Gly					
	80		85		90
Lys Asp Gln Lys Glu Ala Ala Leu Val Asp Met Val Asn Asp Gly					
	95		100		105
Val Glu Asp Leu Arg Cys Lys Tyr Gly Thr Leu Ile Tyr Thr Asn					
	110		115		120
Tyr Glu Asn Gly Lys Asp Asp Tyr Val Lys Ala Leu Pro Gly His					
	125		130		135
Leu Lys Pro Phe Glu Thr Leu Leu Ser Gln Asn Gln Gly Gly Lys					
	140		145		150
Ala Phe Ile Val Gly Asn Gln Ile Ser Phe Ala Asp Tyr Asn Leu					
	155		160		165
Leu Asp Leu Leu Leu Val His Gln Val Leu Ala Pro Gly Cys Leu					
	170		175		180
Asp Asn Phe Arg Pro Ala Leu Cys Leu Cys Gly Ser Pro Gln Cys					
	185		190		195
Pro Pro Gln Asp Lys Ala Phe Leu Ser Ser Pro Asp His Leu Asn					
	200		205		210
Arg Pro Ile Asn					

&lt;210&gt; 285

&lt;211&gt; 212

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:453570.1.orf3:2000SEP08

&lt;400&gt; 285

Cys Thr Leu Leu Arg Leu Ile Pro Ser Ala Ser His Phe Lys Arg					
1	5		10		15
Phe Asp Arg Val Arg Arg Phe Ala Pro Ala Ala Met Ala Thr Ser					
	20		25		30
Ser Gly Pro Lys Glu Ala Pro Ala Asn Asn Pro Gly Leu Gln Thr					
	35		40		45
Glu Val Asp Pro Ala Thr Lys Gly Tyr Phe Leu Gln Gln Thr Met					
	50		55		60
Leu Arg Val Lys Asp Pro Lys Val Ser Leu Asp Phe Tyr Ser Arg					
	65		70		75
Val Met Gly Met Ser Leu Leu Lys Arg Leu Asp Phe Glu Glu Met					
	80		85		90
Lys Phe Ser Leu Tyr Phe Leu Gly Tyr Glu Asp Val Thr Leu Ala					
	95		100		105
Pro Asp Asp His Ile Lys Arg Thr Glu Trp Thr Phe Arg Gln Lys					
	110		115		120
Ala Thr Leu Glu Leu Thr His Asn Trp Gly Thr Glu Asn Asp Pro					
	125		130		135
Glu Phe Lys Gly Tyr His Asn Gly Asn Ser Asp Pro Arg Gly Phe					
	140		145		150
Gly His Ile Gly Val Thr Val Asp Asp Val His Lys Ala Cys Glu					
	155		160		165
Arg Phe Glu Arg Leu Gly Val Glu Phe Val Lys Lys Pro Asp Asp					
	170		175		180
Gly Lys Ile Lys Gly Ile Ala Phe Ile Lys Asp Pro Asp Gly Tyr					
	185		190		195
Trp Ile Glu Ile Phe Asp Gln Thr Ile Gly Thr Val Thr Ser Ser					

Ala Ser 200 205 210

<210> 286  
 <211> 183  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:072072.1.orf2:2000SEP08

<400> 286  
 Leu Leu Leu Pro Gln Val Leu Pro Arg Arg Arg Gly Ser Trp Thr  
 1 5 10 15  
 Ser Arg Pro Pro Gly Ala Ser Glu Ser Pro Ser Pro Gly Leu Pro  
 20 25 30  
 Arg Ser Pro Glu Arg Trp Gly Leu Arg Gly Phe Gly Ser Arg Gln  
 35 40 45  
 Arg Asn Met Ala Pro Trp Thr Leu Trp Arg Cys Cys Gln Arg Cys  
 50 55 60  
 Arg Val Ala Gly Cys Arg Val Leu Phe Ile Thr Phe Val Val Val  
 65 70 75  
 Trp Ser Tyr Tyr Ala Tyr Ala Gly Gly Ala Leu Leu Cys Leu Leu  
 80 85 90  
 Phe Leu Glu Ser Glu Ala Lys Trp Lys Asp Arg Cys Leu Pro Cys  
 95 100 105  
 Gly Phe Pro Ser Val Leu Cys Tyr Val Cys Met Val Leu Leu Asp  
 110 115 120  
 Asp Asn Phe His Ile Ser Arg Ile Pro Pro Pro Lys Ser Ser Thr  
 125 130 135  
 Cys Pro Leu Leu Lys Arg Asn Val Met Lys Lys Glu Phe Ser Pro  
 140 145 150  
 Arg Lys Thr Thr Arg Asn Phe Arg Glu Glu Gln Gln Glu Leu Tyr  
 155 160 165  
 Leu Ser Ile Pro His Gln Leu Gln Lys Leu Ser Asp Ile Val Lys  
 170 175 180  
 Asn Val Ser

<210> 287  
 <211> 164  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:148565.4.orf1:2000SEP08

<400> 287  
 Ala Leu Lys Pro Ser Pro Pro Ser Thr Ser Ala Pro Ile Ser Ala  
 1 5 10 15  
 Pro Leu Pro Leu Gln His Leu Pro Ser Pro Ser Ser Arg Leu Arg  
 20 25 30  
 Leu Leu His Leu Ala Ser His Gly Gln Asp Gln Asp Arg Asn  
 35 40 45  
 Gln Arg Phe Arg Lys Asp Arg Gln Asp Ser Trp Pro Gly Ser Pro  
 50 55 60  
 Cys Arg Ala Arg Met Ser Ser Ser Ser Pro Leu Asn Asp Pro Phe  
 65 70 75  
 Ile Thr Thr Asp Tyr Met Thr Tyr Met Phe Lys Leu Arg His Arg  
 80 85 90

Ala	Arg	Pro	Met	Glu	Ala	Gln	Arg	His	His	Pro	Gln	Gly	Thr	Pro
				95					100					105
Arg	Arg	Phe	Ser	Ser	Ala	Arg	Ser	Arg	Ile	Thr	Val	Ile	Trp	His
				110					115					120
Gln	Glu	Pro	Arg	Gly	Asp	Pro	Val	Gly	Met	Arg	Leu	Ala	Ala	Glu
				125					130					135
Tyr	Val	Val	Glu	Ser	Thr	Gly	Val	Phe	Thr	Asp	Lys	Asp	Lys	Ala
				140					145					150
Ala	Asp	Thr	Ser	Glu	Gly	Trp	Cys	Gln	Glu	Gly	Cys	Tyr	Leu	
				155					160					

&lt;210&gt; 288

&lt;211&gt; 154

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:368626.4.orf1:2000SEP08

&lt;400&gt; 288

Ser	Ser	Ser	Leu	Gln	Ser	Ser	Ser	Ile	Ala	Leu	Pro	Ser	Thr	Leu
1				5					10					15
Leu	Ser	Thr	Val	Gly	Ser	Met	Thr	Leu	Gly	Pro	Gly	Cys	Pro	Met
				20					25					30
Leu	His	His	Pro	Leu	Gly	Lys	Pro	Pro	Pro	Gln	Thr	Lys	Gly	Thr
				35					40					45
Thr	Thr	Leu	Lys	Thr	Tyr	Leu	Asp	Thr	Leu	Pro	Glu	Val	Asn	Ile
				50					55					60
Ser	Cys	Asn	Asn	Leu	Leu	Phe	Trp	Leu	Val	Ser	Gln	Glu	Pro	
				65					70					75
Lys	Asp	Gln	Arg	Pro	Leu	Gly	Thr	Tyr	Pro	Asp	Glu	His	Phe	Thr
				80					85					90
Glu	Glu	Ala	Pro	Arg	Arg	Ser	Ile	Ala	Ala	Phe	Gln	Ser	Arg	Leu
				95					100					105
Ala	Gln	Ile	Ser	Arg	Asp	Ile	Gln	Glu	Arg	Glu	Pro	Arg	Val	Leu
				110					115					120
His	Cys	Leu	Lys	Leu	Thr	Trp	Asn	Leu	Pro	Phe	Ile	Glu	Asn	Ser
				125					130					135
Val	Leu	His	Leu	Asn	His	Pro	Gln	Ile	Pro	Pro	Lys	Lys	Lys	Glu
				140					145					150
Lys	Val	Gln	Ala											

&lt;210&gt; 289

&lt;211&gt; 148

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:346123.1.orf2:2000SEP08

&lt;400&gt; 289

Val	Glu	Gln	Ser	Arg	Ser	Thr	Pro	Ser	Pro	Ile	Ser	Leu	Ser	Ser
1				5					10					15
Pro	Ser	Glu	Ala	Pro	His	Pro	Lys	Ser	Asn	Arg	Thr	Pro	His	Arg
				20					25					30
Asp	Gly	Pro	Gln	Val	Leu	Arg	Arg	Trp	Gln	Leu	Gly	Asn	Ala	Met
				35					40					45
Glu	Pro	Gln	Ile	Arg	Ser	Lys	Lys	Ile	Val	Lys	Thr	Leu	Asn	Glu
				50					55					60
Gly	Gln	Val	Pro	Pro	Ser	Asp	Val	Val	Glu	Val	Val	Cys	Gln	Pro



	65		70		75									
Ser	Leu	Cys	Leu	Pro	Ser	Cys	Gly	Gln	Glu	Pro	Ala	Ala	Pro	Arg
	80								85					90
Val	Pro	Cys	Cys	Cys	Ser	Glu	Leu	Leu	Gly	Glu	Asp	Arg	Glu	Val
	95								100					105
Leu	Ser	Leu	Gly	Glu	Val	Ser	Ala	Glu	Met	Leu	Val	Asn	Leu	Gly
	110								115					120
Gly	Ser	Leu	Gly	His	Ser	Trp	Thr	Leu	Cys	Lys	Gly	Glu	Leu	Cys
	125								130					135
Trp	Glu	Asn	Gln	Met	Asn	Leu	Leu	Glu	Thr	Thr	Val	Ala		
	140								145					

&lt;210&gt; 290

&lt;211&gt; 291

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:335795.11.orf2:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 209

&lt;223&gt; unknown or other

&lt;400&gt; 290

Arg	Lys	Thr	Ile	Ser	Ser	Tyr	Leu	Pro	Phe	Met	Asp	Lys	Arg	Gly
1				5					10					15
Asp	Thr	Ser	Asn	Ile	Ile	Thr	Val	Arg	Val	Ala	Asp	Gly	Gln	Thr
			20						25					30
Val	Gln	Gly	Glu	Val	Trp	Thr	Thr	Thr	Pro	Tyr	Thr	Val	Ala	Ala
			35						40					45
Glu	Ile	Ser	Gln	Glu	Leu	Ala	Glu	Ser	Thr	Val	Ile	Ala	Arg	Val
			50						55					60
Asn	Gly	Glu	Leu	Trp	Asp	Leu	Asp	Arg	Pro	Leu	Glu	Gly	Asp	Ser
			65						70					75
Ser	Leu	Glu	Leu	Leu	Thr	Phe	Asp	Asn	Glu	Glu	Ala	Gln	Ala	Val
			80						85					90
Tyr	Trp	His	Ser	Ser	Ala	His	Ile	Leu	Gly	Glu	Ala	Met	Glu	Leu
			95						100					105
Tyr	Tyr	Gly	Gly	His	Leu	Cys	Tyr	Gly	Pro	Pro	Ile	Glu	Asn	Gly
			110						115					120
Phe	Tyr	Tyr	Asp	Met	Phe	Ile	Glu	Asp	Arg	Ala	Val	Ser	Ser	Thr
			125						130					135
Glu	Leu	Ser	Ala	Leu	Glu	Asn	Ile	Cys	Lys	Ala	Ile	Ile	Lys	Glu
			140						145					150
Lys	Gln	Pro	Phe	Glu	Arg	Leu	Glu	Val	Ser	Lys	Glu	Ile	Leu	Leu
			155						160					165
Glu	Met	Phe	Lys	Tyr	Asn	Lys	Phe	Lys	Cys	Arg	Ile	Leu	Asn	Glu
			170						175					180
Lys	Val	Asn	Thr	Ala	Thr	Thr	Thr	Val	Tyr	Arg	Cys	Gly	Pro	Leu
			185						190					195
Ile	Asp	Leu	Cys	Lys	Gly	Pro	His	Val	Arg	His	Thr	Gly	Xaa	Ile
			200						205					210
Lys	Thr	Ile	Lys	Ile	Phe	Lys	Asn	Ser	Ser	Thr	Tyr	Trp	Glu	Gly
			215						220					225
Asn	Pro	Glu	Met	Glu	Thr	Leu	Gln	Arg	Ile	Tyr	Gly	Ile	Ser	Phe
			230						235					240
Pro	Asp	Asn	Lys	Met	Met	Arg	Asp	Trp	Glu	Lys	Phe	Gln	Glu	Glu
			245						250					255
Ala	Lys	Asn	Arg	Asp	His	Arg	Lys	Ile	Trp	Glu	Gly	Thr	Arg	Thr
			260						265					270

Phe Leu Phe Pro Arg Phe Glu Ser Trp Lys Leu Phe Phe Pro Ser  
 275 280 285  
 Gln Arg Ser Leu His Leu  
 290

<210> 291  
 <211> 354  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:246023.2.orf1:2000SEP08

<400> 291  
 Gly Gly Leu Ala Arg Thr Trp Gln Gly Leu Phe Thr Met Ala Asp  
 1 5 10 15  
 Asp Leu Glu Gln Gln Ser Gln Gly Trp Leu Ser Ser Trp Leu Pro  
 20 25 30  
 Thr Trp Arg Pro Thr Ser Met Ser Gln Leu Lys Asn Val Glu Ala  
 35 40 45  
 Arg Ile Leu Gln Cys Leu Gln Asn Lys Phe Leu Ala Arg Tyr Val  
 50 55 60  
 Ser Leu Pro Asn Gln Asn Lys Ile Trp Thr Val Thr Val Ser Pro  
 65 70 75  
 Glu Gln Asn Asp Arg Thr Pro Leu Val Met Val His Gly Phe Gly  
 80 85 90  
 Gly Gly Val Gly Leu Trp Ile Leu Asn Met Asp Ser Leu Ser Ala  
 95 100 105  
 Arg Arg Thr Leu His Thr Phe Asp Leu Leu Gly Phe Gly Arg Ser  
 110 115 120  
 Ser Arg Pro Ala Phe Pro Arg Asp Pro Glu Gly Ala Glu Asp Glu  
 125 130 135  
 Phe Val Thr Ser Ile Glu Thr Trp Arg Glu Thr Met Gly Ile Pro  
 140 145 150  
 Ser Met Ile Leu Leu Gly His Ser Leu Gly Gly Phe Leu Ala Thr  
 155 160 165  
 Ser Tyr Ser Ile Lys Tyr Pro Asp Arg Val Lys His Leu Ile Leu  
 170 175 180  
 Val Asp Pro Trp Gly Phe Pro Leu Arg Pro Thr Asn Pro Ser Glu  
 185 190 195  
 Ile Arg Ala Pro Pro Ala Trp Val Lys Ala Val Ala Ser Val Leu  
 200 205 210  
 Gly Arg Ser Asn Pro Leu Ala Val Leu Arg Val Ala Gly Pro Trp  
 215 220 225  
 Gly Pro Gly Leu Val Gln Arg Phe Arg Pro Asp Phe Lys Arg Lys  
 230 235 240  
 Phe Ala Asp Phe Phe Glu Asp Asp Thr Ile Ser Glu Tyr Ile Tyr  
 245 250 255  
 His Cys Asn Ala Gln Asn Pro Ser Gly Glu Thr Ala Phe Lys Ala  
 260 265 270  
 Met Met Glu Ser Phe Gly Trp Ala Arg Arg Pro Met Leu Glu Arg  
 275 280 285  
 Ile His Leu Ile Arg Lys Asp Val Pro Ile Thr Met Ile Tyr Gly  
 290 295 300  
 Ser Asp Thr Trp Ile Asp Thr Ser Thr Gly Lys Lys Val Lys Met  
 305 310 315  
 Gln Arg Pro Asp Ser Tyr Val Arg Asp Met Glu Val Lys Gly Ala  
 320 325 330  
 Ser His His Val Tyr Ala Asp Gln Pro His Ile Phe Asn Ala Val  
 335 340 345  
 Val Glu Glu Ile Cys Asp Ser Val Asp  
 350

<210> 292  
 <211> 132  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1100661.1.orf3:2000SEP08

<400> 292  
 Leu Gly Leu Asn Thr Asn Ala Met Ala Cys Pro Ser Leu Ala Cys  
   1                  5                  10                  15  
 Cys Leu Leu Gly Leu Leu Ala Leu Thr Ser Ala Cys Tyr Ile Gln  
                   20                  25                  30  
 Asn Cys Pro Leu Gly Gly Lys Arg Ala Ala Leu Asp Leu Asp Met  
                   35                  40                  45  
 Arg Lys Cys Leu Pro Cys Gly Pro Gly Gly Lys Gly Arg Cys Phe  
                   50                  55                  60  
 Gly Pro Ser Ile Cys Cys Ala Asp Glu Leu Gly Cys Phe Val Gly  
                   65                  70                  75  
 Thr Ala Glu Ala Leu Arg Cys Gln Glu Glu Asn Tyr Leu Pro Ser  
                   80                  85                  90  
 Pro Cys Gln Ser Gly Gln Lys Pro Cys Gly Ser Gly Gly Arg Cys  
                   95                  100                 105  
 Ala Thr Ala Gly Ile Cys Cys Ser Pro Asp Gly Cys Arg Thr Asp  
                  110                 115                 120  
 Pro Ala Cys Asp Pro Glu Ser Ala Phe Ser Glu Arg  
                  125                 130

<210> 293  
 <211> 100  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:475856.1.orf1:2000SEP08

<400> 293  
 Ile Thr Leu Ala Gly Ser Arg Tyr Met Lys Glu Pro Thr Leu Glu  
   1                  5                  10                 15  
 Arg Asn Leu Met Asn Val Lys Gln Cys Gly Lys Ala Phe Ser Ser  
                   20                  25                 30  
 Ser Ser Tyr Ile His Ile His Glu Arg Ile His Thr Gly Glu Lys  
                   35                  40                 45  
 Pro Tyr Glu Cys Lys Glu Cys Gly Lys Pro Phe Ser Phe Leu Thr  
                   50                  55                 60  
 Gly Phe Arg Val His Met Arg Met His Thr Gly Glu Lys Pro Tyr  
                   65                  70                 75  
 Lys Cys Lys Asp Cys Gly Asn Ala Phe Ile Trp Arg Ala Ser Leu  
                   80                  85                 90  
 Gln Tyr His Val Lys Lys Val His Ala Glu  
                   95                 100

<210> 294  
 <211> 146  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1015343.1.orf1:2000SEP08

&lt;400&gt; 294

Gly	Asp	Ala	Thr	Gly	Val	Val	Ser	Ala	Ala	Cys	Gly	Pro	Ala	Ser
1				5					10					15
Arg	Leu	Thr	His	Arg	Ala	Gln	Ala	Arg	Leu	Ser	Glu	Ala	Gly	Glu
				20					25					30
Met	Leu	Ser	Cys	Arg	Leu	Gln	Cys	Ala	Leu	Ala	Ala	Leu	Cys	Ile
				35					40					45
Val	Leu	Ala	Leu	Gly	Gly	Val	Thr	Gly	Ala	Pro	Ser	Asp	Pro	Arg
				50					55					60
Leu	Arg	Gln	Phe	Leu	Gln	Lys	Ser	Leu	Ala	Ala	Ala	Thr	Gly	Lys
				65					70					75
Gln	Glu	Leu	Ala	Lys	Tyr	Phe	Leu	Ala	Glu	Leu	Leu	Ser	Glu	Pro
				80					85					90
Asn	Gln	Thr	Glu	Asn	Asp	Ala	Leu	Glu	Pro	Glu	Asp	Leu	Pro	Gln
				95					100					105
Ala	Ala	Glu	Gln	Asp	Glu	Met	Arg	Leu	Glu	Leu	Gln	Arg	Ser	Ala
				110					115					120
Asn	Ser	Asn	Pro	Ala	Met	Ala	Pro	Arg	Glu	Arg	Lys	Ala	Gly	Cys
				125					130					135
Lys	Asn	Phe	Phe	Trp	Lys	Thr	Phe	Thr	Ser	Cys				
				140					145					

&lt;210&gt; 295

&lt;211&gt; 285

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1400575.1.orf2:2000SEP08

&lt;400&gt; 295

Thr	Leu	Gly	Thr	Thr	Pro	Gly	His	Pro	Arg	Gly	Arg	Lys	Met	Asp
1				5					10					15
Ser	Val	Ala	Phe	Glu	Asp	Val	Ser	Val	Ser	Phe	Ser	Gln	Glu	Glu
				20					25					30
Trp	Ala	Leu	Leu	Ala	Pro	Ser	Gln	Lys	Lys	Leu	Tyr	Arg	Asp	Val
				35					40					45
Met	Gln	Glu	Thr	Phe	Lys	Asn	Leu	Ala	Ser	Ile	Gly	Glu	Lys	Trp
				50					55					60
Glu	Asp	Pro	Asn	Val	Glu	Asp	Gln	His	Lys	Asn	Gln	Gly	Arg	Asn
				65					70					75
Leu	Arg	Ser	His	Thr	Gly	Glu	Arg	Leu	Cys	Glu	Gly	Lys	Glu	Gly
				80					85					90
Ser	Gln	Cys	Ala	Glu	Asn	Phe	Ser	Pro	Asn	Leu	Ser	Val	Thr	Lys
				95					100					105
Lys	Thr	Ala	Gly	Val	Lys	Pro	Tyr	Glu	Cys	Thr	Ile	Cys	Gly	Lys
				110					115					120
Ala	Phe	Met	Arg	Leu	Ser	Ser	Leu	Thr	Arg	His	Met	Arg	Ser	His
				125					130					135
Thr	Gly	Tyr	Glu	Leu	Phe	Glu	Lys	Pro	Tyr	Lys	Cys	Lys	Glu	Cys
				140					145					150
Glu	Lys	Ala	Phe	Ser	Tyr	Leu	Lys	Ser	Phe	Gln	Arg	His	Glu	Arg
				155					160					165
Ser	His	Thr	Gly	Glu	Lys	Pro	Tyr	Lys	Cys	Lys	Gln	Cys	Gly	Lys
				170					175					180
Thr	Phe	Ile	Tyr	His	Gln	Pro	Phe	Gln	Arg	His	Glu	Arg	Thr	His
				185					190					195
Ile	Gly	Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Gln	Cys	Gly	Lys	Ala	Leu
				200					205					210
Ser	Cys	Ser	Ser	Ser	Leu	Arg	Val	His	Glu	Arg	Ile	His	Thr	Gly
				215					220					225
Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Gln	Cys	Gly	Lys	Ala	Phe	Ser	Cys

	230		235		240
Ser Ser Ser Ile	Arg Val His Glu Arg	Thr His Thr Gly Glu Lys			
	245		250		255
Pro Tyr Ala Cys	Lys Glu Cys Gly Lys	Ala Phe Ile Ser His Thr			
	260		265		270
Ser Val Leu Thr	His Met Ile Thr His	Asn Gly Ala Val Pro Ala			
	275		280		285

<210> 296  
 <211> 213  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1080545.1.orf2:2000SEP08

<400> 296

Ile Ser Leu Ser Phe Lys Tyr Met Lys Gly Ser Thr Glu Lys Arg		
1	5	10
Thr Val Cys Asn Val Arg Val Met Gly Lys His Ser Val Cys Pro		
	20	25
Val Tyr Phe Ile Asp Met Lys Gly Leu Thr Leu Glu Glu Lys Pro		
	35	40
Met Asn Cys Lys Gln Cys Gly Arg Ser Phe Asn Cys Ser Ser Ser		
	50	55
Phe Arg Tyr His Gly Arg Thr His Thr Gly Glu Lys Pro Tyr Glu		
	65	70
Cys Lys Gln Cys Gly Lys Ala Phe Arg Ser Ala Ser Gln Leu Gln		
	80	85
Ile His Gly Arg Thr His Thr Gly Glu Lys Pro Tyr Glu Cys Lys		
	95	100
Gln Cys Gly Lys Ala Phe Gly Ser Ala Ser His Leu Gln Met His		
	110	115
Gly Arg Thr His Thr Gly Glu Lys Pro Tyr Glu Cys Lys Gln Cys		
	125	130
Gly Lys Ser Phe Gly Cys Ala Ser Arg Leu Gln Met His Gly Arg		
	140	145
Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Lys Gln Cys Gly Lys		
	155	160
Ala Phe Gly Cys Pro Ser Asn Leu Arg Arg His Gly Arg Thr His		
	170	175
Thr Gly Glu Lys Pro Tyr Lys Cys Asn Gln Cys Gly Lys Val Phe		
	185	190
Arg Cys Ser Ser Gln Leu Gln Val His Gly Arg Ala His Cys Ile		
	200	205
Asp Thr Pro		210

<210> 297  
 <211> 95  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:213947.1.orf1:2000SEP08

<400> 297

Asn Ile Gly Asn Ile Phe Ile Glu Met Phe Val Gly Ile Asn Ile		
1	5	10
Phe Phe Leu Val Val Ala Thr Arg Gly Leu Val Leu Gly Met Leu		

	20		25		30
Gly Asn Gly Leu Ile	Gly Leu Val Asn Cys Ile Glu Trp Ala Lys				
	35		40		45
Ser Trp Lys Val Ser	Ser Ala Asp Phe Ile Leu Thr Ser Leu Ala				
	50		55		60
Ile Val Arg Ile Ile	Arg Leu Tyr Leu Ile Leu Phe Asp Ser Phe				
	65		70		75
Ile Met Val Leu Ser	Pro His Leu Tyr Thr Ile Arg Lys Leu Val				
	80		85		90
Lys Leu Phe Thr Ile					
	95				

&lt;210&gt; 298

&lt;211&gt; 328

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:720641.1.orf2:2000SEP08

&lt;400&gt; 298

Glu Ala Leu Tyr Ser	Asp Leu Arg Leu Val Thr Lys Ser Pro Ser				
1	5		10		15
Ala Pro Leu Ile Asp	Ser Gly Arg Ala His Gln Glu Asp Gly Gln				
	20		25		30
Ser His Cys Ser Glu	His Ile Pro Ser Val Gly Ile Phe Gln Phe				
	35		40		45
Phe Arg Pro Ala Glu	Ser Thr Phe Cys Gly Asp Ser Leu Leu His				
	50		55		60
Val Thr Ile Leu Ala	Ala Asn Val Ser Ile Met Gly Ala Ile Lys				
	65		70		75
Leu Ser His Asn Leu	His Thr Pro Met Tyr Phe Phe Leu Cys Gly				
	80		85		90
Leu Ser Phe Ser Glu	Thr Cys Thr Thr Val Val Val Ile Pro Arg				
	95		100		105
Met Leu Val Asp Phe	Leu Ser Glu Ser Lys Thr Ile Ser Leu Pro				
	110		115		120
Glu Cys Ala Thr Gln	Met Phe Phe Phe Leu Gly Phe Ala Ser Asn				
	125		130		135
Asn Cys Phe Ile Met	Ala Ala Met Ser Tyr Asp Arg Tyr Thr Ala				
	140		145		150
Ile His Asn Pro Leu	Gln Tyr His Thr Leu Met Thr Arg Lys Ile				
	155		160		165
Cys Leu Gln Met Met	Met Ala Ser Trp Met Val Gly Phe Leu Phe				
	170		175		180
Ser Leu Cys Ile Ile	Val Thr Val Phe Asn Leu Ser Leu Cys Asp				
	185		190		195
Leu Asn Thr Ile Gln	His Tyr Phe Cys Asp Ile Ser Pro Val Val				
	200		205		210
Ser Leu Ala Cys Asn	Tyr Thr Phe Tyr His Glu Met Ala Ile Phe				
	215		220		225
Val Leu Ser Ala Phe	Val Leu Val Gly Ser Cys Ile Leu Ile Met				
	230		235		240
Ile Ser Tyr Val Phe	Ile Val Phe Ile Val Ile Lys Met Pro Ser				
	245		250		255
Ala Lys Gly Arg Ser	Lys Ala Phe Ser Thr Cys Ser Ser His Leu				
	260		265		270
Thr Val Val Ser Ile	His Tyr Gly Phe Ala Cys Phe Val Tyr Leu				
	275		280		285
Arg Pro Lys Asn Ser	Asn Ser Phe Asp Glu Asp Met Leu Thr Ala				
	290		295		300
Met Ile Tyr Thr Ile	Leu Met Pro Leu Leu Asn Pro Ile Val Tyr				



&lt;223&gt; Incyte ID No: LI:1178118.1.orf3:2000SEP08

&lt;400&gt; 301

Arg	Thr	Cys	Pro	Asp	Ser	Leu	Trp	Leu	Pro	Leu	Leu	Arg	Ile	Leu
1				5					10					15
Leu	Val	Gln	Thr	Gln	Lys	Thr	Val	His	Glu	Ala	Arg	Asp	Ile	Arg
				20					25					30
His	Ala	Pro	Asn	Asn	Leu	Ala	Ser	Val	Ala	Phe	Pro	Phe	Pro	Lys
				35					40					45
Pro	Ala	Leu	Ile	Ser	His	Leu	Glu	Arg	Gly	Glu	Ala	Pro	Trp	Gly
				50					55					60
Pro	Asp	Pro	Trp	Asp	Thr	Glu	Ile	Leu	Arg	Gly	Ile	Ser	Gln	Gly
				65					70					75
Gly	Glu	Ser	Trp	Ile	Lys	Asn	Glu	Gly	Leu	Val	Ile	Lys	Gln	Glu
				80					85					90
Ala	Ser	Glu	Glu	Thr	Glu	Leu	His	Arg	Met	Pro	Val	Gly	Gly	Leu
				95					100					105
Leu	Arg	Asn	Val	Ser	Gln	His	Phe	Asp	Phe	Lys	Arg	Lys	Ala	Leu
				110					115					120
Lys	Gln	Thr	Phe	Asn	Leu	Asn	Pro	Asn	Leu	Ile	Leu	Arg	Gly	Gly
				125					130					135
Met	Lys	Phe	Tyr	Glu	Cys	Lys	Glu	Cys	Gly	Lys	Ile	Phe	Arg	Tyr
				140					145					150
Asn	Ser	Lys	Leu	Ile	Arg	His	Gln	Met	Ser	His	Thr	Gly	Glu	Lys
				155					160					165
Pro	Phe	Lys	Cys	Lys	Glu	Cys	Gly	Lys	Ala	Phe	Lys	Ser	Ser	Tyr
				170					175					180
Asp	Cys	Ile	Val	His	Glu	Lys	Asn	His	Ile	Gly	Glu	Gly	Pro	Tyr
				185					190					195
Glu	Cys	Lys	Glu	Cys	Gly	Lys	Gly	Leu	Ser	Ser	Asn	Thr	Ala	Leu
				200					205					210
Thr	Gln	His	Gln	Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Tyr	Glu	Cys
				215					220					225
Lys	Glu	Cys	Gly	Lys	Ala	Phe	Arg	Arg	Ser	Ala	Ala	Tyr	Leu	Gln
				230					235					240
His	Gln	Arg	Leu	His	Thr	Gly	Glu	Lys	Leu	Tyr	Lys	Cys	Lys	Glu
				245					250					255
Cys	Trp	Lys	Ala	Phe	Gly	Cys	Arg	Ser	Leu	Phe	Ile	Val	His	Gln
				260					265					270
Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Tyr	Gln	Cys	Lys	Glu	Cys	Gly
				275					280					285
Lys	Ala	Phe	Thr	Gln	Lys	Ile	Ala	Ser	Ile	Gln	His	Gln	Arg	Val
				290					295					300
His	Thr	Gly	Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Val	Cys	Gly	Lys	Ala
				305					310					315
Phe	Lys	Trp	Tyr	Gly	Ser	Phe	Val	Gln	His	Gln	Lys	Leu	His	Pro
				320					325					330
Val	Glu	Lys	Lys	Pro	Val	Lys	Val	Leu	Gly	Pro	Ser	Leu	Val	Ser
				335					340					345
Pro	Gln	Cys	Ser	Ser	Pro	Ala	Ile	Pro	Pro	Val	Leu	Leu	Gln	Gly
				350					355					360
Ser	Cys	Ser	Ala	Ser	Ala	Val	Ala	Val	Pro	Ser	Leu	Thr	Phe	Pro
				365					370					375
His	Ala	Val	Leu	Ile	Pro	Thr	Ser	Gly	Asn	Phe	Phe	Met	Leu	Leu
				380					385					390
Pro	Thr	Ser	Gly	Ile	Pro	Ser	Ser	Ser	Ala	Gln	Ile	Val	Arg	Val
				395					400					405
Phe	Gln	Gly	Leu	Thr	Pro	Thr	Val	Lys	Pro	Ser	Pro	Val	Ile	Leu
				410					415					420
Thr	Pro	Ser	Ser	His	Ser	Ser								
				425										

&lt;210&gt; 302



<211> 95  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:213947.1.orf3:2000SEP08

<220>  
 <221> unsure  
 <222> 1-2  
 <223> unknown or other

<400> 302  
 Xaa Xaa Gly Asn Ile Phe Ile Glu Met Phe Val Gly Ile Asn Ile  
 1 5 10 15  
 Phe Phe Leu Val Val Ala Thr Arg Gly Leu Val Leu Gly Met Leu  
 20 25 30  
 Gly Asn Gly Leu Ile Gly Leu Val Asn Cys Ile Glu Trp Ala Lys  
 35 40 45  
 Ser Trp Lys Val Ser Ser Ala Asp Phe Ile Leu Thr Ser Leu Ala  
 50 55 60  
 Ile Val Arg Ile Ile Arg Leu Tyr Leu Ile Leu Phe Asp Ser Phe  
 65 70 75  
 Ile Met Val Leu Ser Pro His Leu Tyr Thr Ile Arg Lys Leu Val  
 80 85 90  
 Lys Leu Phe Thr Ile  
 95

<210> 303  
 <211> 82  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:407304.1.orf1:2000SEP08

<400> 303  
 Ile Leu Ser Gln Thr Ser Lys Lys Met Arg Val Ile Ala Ser Gln  
 1 5 10 15  
 Ser Ala Leu Gln Asn Met Leu Lys Ala Ala Leu Lys Gly Glu Gly  
 20 25 30  
 Lys Arg Tyr Arg Leu Glu Thr Gln Ile His Ile Lys Glu Leu Ser  
 35 40 45  
 Gly Lys Glu Tyr Met Lys Thr Ile Glu Asn Leu Leu His Phe Leu  
 50 55 60  
 Phe Leu Val Asp Leu Val Gly Asn Cys Leu Ala Lys Val Ile Ile  
 65 70 75  
 Val Thr Leu Cys Trp Gly Pro  
 80

<210> 304  
 <211> 407  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:337358.1.orf2:2000SEP08

<400> 304  
 Glu Leu Leu Leu His His Trp Ser Arg Pro Tyr Leu Ser Ala Pro

1	5	10	15
Ser Pro Arg Ala	Pro Gly Gly Asp Pro	Arg Ile Gly Arg Pro	Gln
	20	25	30
Ala Ser Gly His	Leu Ala Gln Gln Glu	Ala Pro Gly Ser Gly	Gln
	35	40	45
Gly Arg Ala Gly	Gly Gly Glu Pro Glu	Pro Ala Pro Cys Pro	Gly
	50	55	60
Pro Ala Glu Pro	Ser Glu Pro Thr His	Gly Ala Pro Ala Pro	Cys
	65	70	75
Ala Ser Leu Pro	Gly Arg Ala Gln Pro	Gly Val Pro Ser Ser	Ala
	80	85	90
Gly Glu Asp Pro	Arg Ala Val Thr Arg	Glu Pro Pro Gln Thr	Leu
	95	100	105
Gly Gly Ser Ala	Ala Gly Ala Ala Gly	Ser Ser Pro Gln Leu	Pro
	110	115	120
Ala Leu Pro Gly	Ser Ser Leu Ser Arg	Ala Arg Gly Pro Ala	Arg
	125	130	135
His Ser Gln Pro	Arg Ala Met Met Lys	Thr Leu Ser Ser Gly	Asn
	140	145	150
Cys Thr Leu Ser	Val Pro Ala Lys Asn	Ser Tyr Arg Met Val	Val
	155	160	165
Leu Gly Ala Ser	Arg Val Gly Lys Ser	Ser Ile Val Ser Arg	Phe
	170	175	180
Leu Asn Gly Arg	Phe Glu Asp Gln Tyr	Thr Pro Thr Ile Glu	Asp
	185	190	195
Phe His Arg Lys	Val Tyr Asn Ile Arg	Gly Asp Met Tyr Gln	Leu
	200	205	210
Asp Ile Leu Asp	Thr Ser Gly Asn His	Pro Phe Pro Ala Met	Arg
	215	220	225
Arg Leu Ser Ile	Leu Thr Gly Asp Val	Phe Ile Leu Val Phe	Ser
	230	235	240
Leu Asp Asn Arg	Glu Ser Phe Asp Glu	Val Lys Arg Leu Gln	Lys
	245	250	255
Gln Ile Leu Glu	Val Lys Ser Cys Leu	Lys Asn Lys Thr Lys	Glu
	260	265	270
Ala Ala Glu Leu	Pro Met Val Ile Cys	Gly Asn Lys Asn Asp	His
	275	280	285
Gly Glu Leu Cys	Arg Gln Val Pro Thr	Thr Glu Ala Glu Leu	Leu
	290	295	300
Val Ser Gly Asp	Glu Asn Cys Ala Tyr	Phe Glu Val Ser Ala	Lys
	305	310	315
Lys Asn Thr Asn	Val Asp Glu Met Phe	Tyr Val Leu Phe Ser	Met
	320	325	330
Ala Lys Leu Pro	His Glu Met Ser Pro	Ala Leu His Arg Lys	Ile
	335	340	345
Ser Val Gln Tyr	Gly Asp Ala Phe His	Pro Arg Pro Phe Cys	Met
	350	355	360
Arg Arg Val Lys	Glu Met Asp Ala Tyr	Gly Met Val Ser Pro	Phe
	365	370	375
Ala Arg Arg Pro	Ser Val Asn Ser Asp	Leu Lys Tyr Ile Lys	Ala
	380	385	390
Lys Val Leu Arg	Glu Gly Gln Ala Arg	Glu Arg Asp Lys Cys	Thr
	395	400	405
Ile Gln			

&lt;210&gt; 305

&lt;211&gt; 277

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:986090.1.orf1:2000SEP08

&lt;400&gt; 305

```

Pro Gln Phe Pro Asp His Thr Val Pro Pro Ser Pro Gly Ala Ala
 1          5          10          15
Pro Arg Gln Glu Glu Val Cys Ser Phe Asn Asn Met Ser Arg Glu
          20          25          30
Glu Asn Val Tyr Met Ala Lys Leu Ala Glu Gln Ala Glu Arg Tyr
          35          40          45
Glu Glu Met Val Glu Tyr Met Glu Lys Val Ala Lys Thr Val Asp
          50          55          60
Val Glu Glu Leu Thr Val Glu Glu Arg Asn Leu Leu Ser Val Ala
          65          70          75
Tyr Lys Asn Val Ile Gly Ala Arg Arg Ala Ser Trp Arg Ile Val
          80          85          90
Ser Ser Ile Glu Gln Lys Glu Glu Ser Arg Lys Asn Glu Glu His
          95          100          105
Val Asn Leu Ile Lys Glu Tyr Arg Gly Lys Ile Glu Ala Glu Leu
          110          115          120
Ser Asn Ile Cys Asp Gly Ile Leu Lys Leu Leu Asp Ser His Leu
          125          130          135
Val Pro Ser Ser Thr Ala Ala Glu Ser Lys Val Phe Tyr Leu Lys
          140          145          150
Met Lys Gly Asp Tyr His Arg Tyr Leu Ala Glu Phe Lys Thr Gly
          155          160          165
Ala Glu Arg Lys Glu Ser Ala Glu Ser Thr Met Val Ala Tyr Lys
          170          175          180
Ala Ala Gln Asp Ile Ala Leu Ala Glu Leu Ala Pro Thr His Pro
          185          190          195
Ile Arg Leu Gly Leu Ala Leu Asn Phe Ser Val Phe Tyr Tyr Glu
          200          205          210
Ile Leu Asn Ser Pro Asp Lys Ala Cys Asn Leu Ala Lys Gln Ala
          215          220          225
Phe Asp Glu Ala Ile Ser Glu Leu Asp Ser Leu Gly Glu Glu Ser
          230          235          240
Tyr Lys Asp Ser Thr Leu Ile Met Gln Leu Leu Arg Asp Asn Leu
          245          250          255
Thr Leu Arg Thr Ser Asp Leu Thr Glu Asp Gly Ala Asp Glu Gly
          260          265          270
Lys Asp Ala Pro Lys Gly Asp
          275

```

&lt;210&gt; 306

&lt;211&gt; 208

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:123250.1.orf2:2000SEP08

&lt;400&gt; 306

```

Lys His Arg Gly Gly Pro Ser Ser Leu Gly Ala Glu Lys Leu Glu
 1          5          10          15
Ser Gly Gln Gly Glu Glu Leu Ser Pro Thr Leu Ala Arg Pro Pro
          20          25          30
Ala Ser Ser Pro His Ala Gly Ser Met Asn Arg Lys Asp Ser Lys
          35          40          45
Arg Lys Ser His Gln Glu Cys Thr Gly Lys Thr Gly Gly Arg Gly
          50          55          60
Arg Pro Arg Gln Val Arg Arg His Lys Thr Cys Pro Ser Pro Arg
          65          70          75
Glu Ile Ser Lys Val Met Ala Ser Met Asn Leu Gly Leu Leu Ser

```

	80		85		90
Glu Gly Gly Cys Ser	Glu Asp Glu Leu Leu	Glu Lys Cys Ile Gln			
	95		100		105
Ser Phe Asp Ser Ala	Gly Ser Leu Cys His	Glu Asp His Met Leu			
	110		115		120
Asn Met Val Leu Ala	Met His Ser Trp Val	Leu Pro Ser Ala Asp			
	125		130		135
Leu Ala Ala Arg Leu	Leu Thr Ser Tyr Gln	Lys Ala Thr Gly Asp			
	140		145		150
Thr Gln Glu Leu Arg	Arg Leu Gln Ile Cys	His Leu Val Arg Tyr			
	155		160		165
Trp Leu Met Arg His	Pro Glu Val Met His	Gln Asp Pro Gln Leu			
	170		175		180
Glu Glu Val Ile Gly	Arg Phe Trp Ala Thr	Val Ala Arg Glu Gly			
	185		190		195
Asn Ser Ala Gln Arg	Arg Leu Gly Asp Ser	Ser Ser Asp Leu			
	200		205		

&lt;210&gt; 307

&lt;211&gt; 87

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1028774.2.orf1:2000SEP08

&lt;400&gt; 307

His Leu Met Asp Ser	Phe Ala Leu Gln Leu	Phe His Met Leu Trp			
1	5	10			15
Leu Ser Ile Gly Gly	Phe Val Ser Met Ala	Gln Leu Thr Met Glu			
	20	25			30
Phe Ser Ala Ser Leu	His Val His Pro Ser	Val Gln Phe Leu Leu			
	35	40			45
His Cys Asp His Ser	Lys Val Gly Ile Pro	Cys Thr Asp Leu Pro			
	50	55			60
Gly Thr Thr Ala Thr	Leu Asp Asn Ile Gln	Phe Leu Gly Gly Gly			
	65	70			75
Phe Ser Thr Thr Pro	Asn Ile Asp Phe Thr	Gly Ser			
	80	85			

&lt;210&gt; 308

&lt;211&gt; 205

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:338927.6.orf2:2000SEP08

&lt;400&gt; 308

Ala Pro Ala Tyr Ser	Ser Gly Pro Gln Glu	Gln Pro Asp Pro Pro			
1	5	10			15
Pro Phe Ser His Pro	Leu Arg Val Ala Pro	Pro Ser Leu Pro Lys			
	20	25			30
Ser Ser Glu Ala Ser	Gly Ser Leu Thr Leu	Gln His His Met Leu			
	35	40			45
Glu Pro Val Gln Arg	Ile Pro Arg Tyr Glu	Leu Leu Leu Lys Glu			
	50	55			60
Tyr Ile Gln Lys Leu	Pro Ala Gln Ala Pro	Asp Gln Ala Asp Ala			
	65	70			75
Gln Lys Ala Leu Asp	Met Ile Phe Ser Ala	Ala Gln His Ser Asn			
	80	85			90

```
<210> 309
<211> 545
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: LG:332944.2.orf1:2000SEP08
```

148/283

Phe	Asp	Gln	Gly	Leu	Asp	Val	Val	Leu	Asp	Asp	Asn	Gln	Asn	Val	
				275					280					285	
His	Asp	Val	Ala	Ala	Leu	Leu	Lys	Glu	Phe	Phe	Arg	Asp	Met	Lys	
				290					295					300	
Asp	Phe	Leu	Leu	Pro	Asp	Asp	Leu	Tyr	Met	Ser	Phe	Leu	Leu	Thr	
				305					310					315	
Ala	Thr	Leu	Lys	Pro	Gln	Asp	Gln	Leu	Ser	Ala	Leu	Gln	Leu	Leu	
				320					325					330	
Val	Tyr	Leu	Met	Pro	Pro	Cys	His	Ser	Asp	Thr	Leu	Glu	Arg	Leu	
				335					340					345	
Leu	Lys	Ala	Leu	His	Lys	Ile	Thr	Glu	Asn	Cys	Glu	Asp	Ser	Ile	
				350					355					360	
Gly	Ile	Asp	Gly	Gln	Leu	Val	Pro	Gly	Asn	Arg	Met	Thr	Ser	Thr	
				365					370					375	
Asn	Leu	Ala	Leu	Val	Phe	Gly	Ser	Ala	Leu	Leu	Lys	Lys	Gly	Lys	
				380					385					390	
Phe	Gly	Lys	Arg	Glu	Ser	Arg	Lys	Thr	Lys	Leu	Gly	Ile	Asp	His	
				395					400					405	
Tyr	Val	Ala	Ser	Val	Asn	Val	Val	Arg	Ala	Met	Ile	Asp	Asn	Trp	
				410					415					420	
Asp	Val	Leu	Phe	Gln	Val	Pro	Pro	His	Ile	Gln	Arg	Gln	Val	Ala	
				425					430					435	
Lys	Arg	Val	Trp	Lys	Ser	Ser	Pro	Glu	Ala	Leu	Asp	Phe	Ile	Arg	
				440					445					450	
Arg	Arg	Asn	Leu	Arg	Lys	Ile	Gln	Ser	Ala	Arg	Ile	Lys	Met	Glu	
				455					460					465	
Glu	Asp	Ala	Leu	Leu	Ser	Asp	Pro	Val	Glu	Thr	Ser	Ala	Glu	Ala	
				470					475					480	
Arg	Ala	Ala	Val	Leu	Ala	Gln	Ser	Lys	Pro	Ser	Asp	Glu	Gly	Ser	
				485					490					495	
Ser	Glu	Glu	Pro	Ala	Val	Pro	Ser	Gly	Thr	Ala	Arg	Ser	His	Asp	
				500					505					510	
Asp	Glu	Glu	Gly	Ala	Gly	Asn	Pro	Pro	Ile	Pro	Glu	Gln	Asp	Arg	
				515					520					525	
Pro	Leu	Leu	Arg	Val	Pro	Arg	Glu	Lys	Glu	Ala	Lys	Thr	Gly	Val	
				530					535					540	
Ser	Tyr	Phe	Phe	Pro											
				545											

&lt;210&gt; 310

&lt;211&gt; 212

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:347174.5.orf2:2000SEP08

&lt;400&gt; 310

Met	Gln	Pro	Pro	Arg	Pro	Ala	Asp	Phe	Lys	Leu	Gln	Val	Ile	Ile	
					5				10					15	
Ile	Gly	Ser	Arg	Gly	Val	Gly	Lys	Thr	Ser	Leu	Met	Glu	Arg	Phe	
					20				25					30	
Thr	Asp	Asp	Thr	Phe	Cys	Glu	Ala	Cys	Lys	Ser	Thr	Val	Gly	Val	
					35				40					45	
Asp	Phe	Lys	Ile	Lys	Thr	Val	Glu	Leu	Arg	Gly	Lys	Lys	Ile	Arg	
					50				55					60	
Leu	Gln	Ile	Trp	Asp	Thr	Ala	Gly	Gln	Glu	Arg	Phe	Asn	Ser	Ile	
					65				70					75	
Thr	Ser	Ala	Tyr	Tyr	Arg	Ser	Ala	Lys	Gly	Ile	Ile	Leu	Val	Tyr	
					80				85					90	
Asp	Ile	Thr	Lys	Lys	Glu	Thr	Phe	Asp	Asp	Leu	Pro	Lys	Trp	Met	
					95				100					105	

Lys	Met	Ile	Asp	Lys	Tyr	Ala	Ser	Glu	Asp	Ala	Glu	Leu	Leu	Leu	
				110					115					120	
Val	Gly	Asn	Lys	Leu	Asp	Cys	Glu	Thr	Asp	Arg	Glu	Ile	Thr	Arg	
				125					130					135	
Gln	Gln	Gly	Glu	Lys	Phe	Ala	Gln	Gln	Ile	Thr	Gly	Met	Arg	Phe	
				140					145					150	
Cys	Glu	Ala	Ser	Ala	Lys	Asp	Asn	Phe	Asn	Val	Asp	Glu	Ile	Phe	
				155					160					165	
Leu	Lys	Leu	Val	Asp	Asp	Ile	Leu	Lys	Lys	Met	Pro	Leu	Asp	Ile	
				170					175					180	
Leu	Arg	Asn	Glu	Leu	Ser	Asn	Ser	Ile	Leu	Ser	Leu	Gln	Pro	Glu	
				185					190					195	
Pro	Glu	Ile	Pro	Pro	Glu	Leu	Pro	Pro	Pro	Arg	Pro	His	Val	Arg	
				200					205					210	

Cys Cys

&lt;210&gt; 311

&lt;211&gt; 189

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:477070.1.orf2:2000SEP08

&lt;400&gt; 311

Lys	Asn	Phe	Ser	Ala	Met	Asn	Lys	Leu	Lys	Lys	Met	Ala	Leu	Arg	
1				5					10					15	
Val	Ile	Ala	Glu	Ser	Leu	Ser	Glu	Glu	Glu	Ile	Gly	Gly	Leu	Lys	
				20					25					30	
Glu	Leu	Phe	Lys	Met	Ile	Asp	Thr	Asp	Asn	Ser	Gly	Thr	Ile	Thr	
				35					40					45	
Tyr	Asp	Glu	Leu	Lys	Asp	Gly	Leu	Lys	Arg	Val	Gly	Ser	Asp	Leu	
				50					55					60	
Met	Glu	Pro	Glu	Ile	Gln	Ala	Leu	Met	Asp	Ala	Ala	Asp	Ile	Asp	
				65					70					75	
Asn	Ser	Gly	Thr	Ile	Asp	Tyr	Gly	Glu	Phe	Leu	Ala	Ala	Thr	Leu	
				80					85					90	
His	Met	Asn	Lys	Leu	Glu	Arg	Glu	Glu	Ser	Leu	Val	Ser	Ala	Phe	
				95					100					105	
Ala	Phe	Phe	Asp	Lys	Asp	Gly	Ser	Gly	Phe	Ile	Thr	Ile	Asp	Glu	
				110					115					120	
Leu	Ser	Gln	Ala	Cys	Glu	Gln	Phe	Gly	Leu	Ser	Asp	Val	His	Leu	
				125					130					135	
Glu	Asp	Met	Ile	Lys	Asp	Val	Asp	Gln	Asn	Asn	Asp	Gly	Gln	Ile	
				140					145					150	
Asp	Tyr	Ser	Glu	Phe	Ala	Ala	Met	Met	Arg	Lys	Gly	Asn	Ala	Gly	
				155					160					165	
Gly	Ala	Gly	Arg	Arg	Thr	Met	Arg	Asn	Ser	Leu	His	Val	Asn	Leu	
				170					175					180	
Gly	Glu	Leu	Leu	Lys	Pro	Thr	Glu	Thr							
				185											

&lt;210&gt; 312

&lt;211&gt; 202

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:723144.1.orf2:2000SEP08

&lt;400&gt; 312

```

Phe Leu Leu Arg Ile Ser His Ser Ala Gln Arg Arg Cys Glu Gly
 1          5          10          15
Arg Gly Thr Arg Arg Lys Met Gly Leu Ala Phe Gly Lys Leu Phe
 20          25          30
Ser Arg Leu Phe Ala Lys Lys Glu Met Arg Ile Leu Met Val Gly
 35          40          45
Leu Asp Ala Ala Gly Lys Thr Thr Ile Leu Tyr Lys Leu Lys Leu
 50          55          60
Gly Glu Ile Val Thr Thr Ile Pro Thr Ile Gly Phe Asn Val Glu
 65          70          75
Thr Val Glu Tyr Lys Asn Ile Ser Phe Thr Val Trp Asp Val Gly
 80          85          90
Gly Gln Asp Lys Ile Arg Pro Leu Trp Arg His Tyr Phe Gln Asn
 95          100         105
Thr Gln Gly Leu Ile Phe Val Val Asp Ser Asn Asp Arg Asp Arg
 110         115         120
Val Val Glu Ala Arg Asp Glu Leu His Arg Met Leu Asn Glu Asp
 125         130         135
Glu Leu Arg Asp Ala Val Leu Leu Val Phe Ala Asn Lys Gln Asp
 140         145         150
Leu Pro Asn Ala Met Asn Ala Ala Glu Ile Thr Asp Lys Leu Gly
 155         160         165
Leu His Ser Leu Arg Gln Arg His Trp Tyr Ile Gln Ser Thr Cys
 170         175         180
Ala Thr Thr Gly Glu Gly Leu Tyr Glu Gly Leu Asp Trp Leu Ser
 185         190         195
Ser Asn Ile Ala Ser Lys Ala
 200

```

&lt;210&gt; 313

&lt;211&gt; 178

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1007188.1.orf1:2000SEP08

&lt;400&gt; 313

```

Glu Glu Phe Cys Ala Phe Tyr Lys Met Met Ser Thr Arg Arg Asp
 1          5          10          15
Leu Tyr Leu Leu Met Leu Thr Tyr Ser Asn His Lys Asp His Leu
 20          25          30
Asp Ala Ser Asp Leu Gln Arg Phe Leu Glu Val Glu Gln Lys Met
 35          40          45
Ser Gly Val Thr Leu Glu Ser Cys Gln Ser Ile Ile Glu Gln Phe
 50          55          60
Glu Pro Cys Leu Glu Asn Lys Ser Lys Gly Val Leu Gly Ile Asp
 65          70          75
Gly Phe Thr Asn Tyr Thr Arg Ser Pro Ala Gly Asp Ile Phe Asn
 80          85          90
Pro Glu His His Arg Val His Gln Asp Met Thr Gln Pro Leu Ser
 95          100         105
His Tyr Phe Ile Thr Ser Ser His Asn Thr Tyr Leu Val Gly Asp
 110         115         120
Gln Leu Met Ser Gln Ser Arg Val Asp Met Tyr Ala Trp Val Leu
 125         130         135
Gln Ala Gly Cys Arg Cys Val Glu Val Asp Cys Trp Asp Gly Pro
 140         145         150
Asp Gly Glu Pro Ile Val His His Gly Tyr Thr Leu Thr Ser Lys
 155         160         165
Ile Leu Phe Lys Asp Val Ile Glu Thr Ile Asn Lys Tyr

```



170

175

<210> 314  
 <211> 77  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1024412.1.orf3:2000SEP08

<400> 314  
 Thr Pro Gly Pro Arg Ser Trp Leu Met Met Ser Gly Thr Asn Asn  
 1 5 10 15  
 Val Ala Gln Ala Arg Lys Leu Val Glu Gln Leu Arg Ile Glu Ala  
 20 25 30  
 Gly Ile Glu Arg Ile Lys Val Ser Lys Ala Ser Ser Glu Leu Met  
 35 40 45  
 Ser Tyr Cys Glu Gln His Ala Arg Asn Asp Pro Leu Leu Val Gly  
 50 55 60  
 Val Pro Ala Ser Glu Asn Pro Phe Lys Asp Lys Lys Pro Cys Ile  
 65 70 75  
 Ile Leu

<210> 315  
 <211> 213  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:284797.3.orf3:2000SEP08

<400> 315  
 Pro Pro Ala Cys Cys Thr Arg Gly Asn Tyr Leu Asn Arg Ser Leu  
 1 5 10 15  
 Ser Ala Gly Ser Asp Ser Glu Gln Leu Ala Asn Ile Ser Val Glu  
 20 25 30  
 Glu Leu Asp Glu Ile Arg Glu Ala Phe Arg Val Leu Asp Arg Asp  
 35 40 45  
 Gly Asn Gly Phe Ile Ser Lys Gln Glu Leu Gly Met Ala Met Arg  
 50 55 60  
 Ser Leu Gly Tyr Met Pro Ser Glu Val Glu Leu Ala Ile Ile Met  
 65 70 75  
 Gln Arg Leu Asp Met Asp Gly Asp Gly Gln Val Asp Phe Asp Glu  
 80 85 90  
 Phe Met Thr Ile Leu Gly Pro Lys Leu Val Ser Ser Glu Gly Arg  
 95 100 105  
 Asp Gly Phe Leu Gly Asn Thr Ile Asp Ser Ile Phe Trp Gln Phe  
 110 115 120  
 Asp Met Gln Arg Ile Thr Leu Glu Glu Leu Lys His Ile Leu Tyr  
 125 130 135  
 His Ala Phe Arg Asp His Leu Thr Met Lys Asp Ile Glu Asn Ile  
 140 145 150  
 Ile Ile Asn Glu Glu Ser Leu Asn Glu Thr Ser Gly Asn Cys  
 155 160 165  
 Gln Thr Glu Phe Glu Gly Val His Ser Gln Lys Gln Asn Arg Gln  
 170 175 180  
 Thr Cys Val Arg Lys Ser Leu Ile Cys Ala Phe Ala Met Ala Phe  
 185 190 195  
 Ile Ile Ser Val Met Leu Ile Ala Ala Asn Gln Ile Leu Arg Ser  
 200 205 210

Gly Met Glu

<210> 316  
 <211> 49  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1092901.1.orf3:2000SEP08

<400> 316  
 Arg Pro Glu Lys Ser Arg Ser Val Ser Ser Met Ser Phe Cys Gln  
           1                  5                  10                  15  
 Cys Ser Lys Leu Leu Pro Leu Leu Phe Leu Phe Tyr Ser His Phe  
                   20                  25                  30  
 Leu Ile Asn Pro Ile Leu Tyr Lys Phe Leu Ile Met Tyr Leu Leu  
                   35                  40                  45  
 Tyr Tyr Phe Leu

<210> 317  
 <211> 235  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:228930.1.orf2:2000SEP08

<400> 317  
 Ala Gln Met Ala Gly Ala Gln Pro Gly Val His Ala Leu Gln Leu  
           1                  5                  10                  15  
 Lys Pro Val Cys Val Ser Asp Ser Leu Lys Lys Gly Thr Lys Phe  
                   20                  25                  30  
 Val Lys Trp Asp Asp Asp Ser Thr Ile Val Thr Pro Ile Ile Leu  
                   35                  40                  45  
 Arg Thr Asp Pro Gln Gly Phe Tyr Phe Tyr Arg Thr Asp Gln Asn  
                   50                  55                  60  
 Lys Glu Thr Glu Leu Leu Asp Leu Ser Leu Val Lys Asp Ala Arg  
                   65                  70                  75  
 Cys Gly Arg His Ala Lys Ala Pro Lys Asp Pro Lys Leu Arg Glu  
                   80                  85                  90  
 Leu Leu Asp Val Gly Asn Ile Gly Arg Leu Glu Gln Arg Met Ile  
                   95                  100                  105  
 Thr Val Val Tyr Gly Pro Asp Leu Val Asn Ile Ser His Leu Asn  
                   110                  115                  120  
 Leu Val Ala Phe Gln Glu Glu Val Ala Lys Glu Trp Thr Asn Glu  
                   125                  130                  135  
 Val Phe Ser Leu Ala Thr Asn Leu Leu Ala Gln Asn Met Ser Arg  
                   140                  145                  150  
 Asp Ala Phe Leu Glu Lys Ala Tyr Thr Lys Leu Lys Leu Gln Val  
                   155                  160                  165  
 Thr Pro Glu Gly Arg Ile Pro Leu Lys Asn Ile Tyr Arg Leu Phe  
                   170                  175                  180  
 Ser Ala Asp Arg Lys Arg Val Glu Thr Ala Leu Glu Ala Cys Ser  
                   185                  190                  195  
 Leu Pro Ser Ser Arg Val Glu Lys Ala Asn Glu Ala Ala Lys Ser  
                   200                  205                  210  
 Glu Gln Ser Cys Gly Lys Ala Pro Pro Lys His Phe His Leu His  
                   215                  220                  225  
 Phe Ile Lys Gln Asn Lys Met Leu Glu Pro

230

235

<210> 318  
 <211> 208  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:722913.1.orf1:2000SEP08

<400> 318  
 Arg Leu Gln Glu Glu Gly Gly Ala Ala Gly Arg Val Gly Glu  
 1 5 10 15  
 Met Phe Leu Trp Asp Trp Phe Tyr Gly Val Leu Ala Ser Leu Gly  
 20 25 30  
 Leu Trp Gln Lys Glu Ala Lys Ile Leu Phe Leu Gly Leu Asp Asn  
 35 40 45  
 Ala Gly Lys Thr Thr Leu Leu His Met Leu Lys Asp Glu Arg Leu  
 50 55 60  
 Val Gln His Gln Pro Thr Gln His Pro Thr Ser Glu Glu Leu Ser  
 65 70 75  
 Ile Gly Lys Ile Lys Phe Lys Ala Phe Asp Leu Gly Gly His Gln  
 80 85 90  
 Ile Ala Arg Arg Val Trp Lys Asp Tyr Tyr Ala Lys Val Asp Ala  
 95 100 105  
 Val Val Tyr Leu Val Asp Ala Tyr Asp Lys Glu Arg Phe Ala Glu  
 110 115 120  
 Ser Lys Lys Glu Leu Asp Ala Leu Leu Ser Asp Asp Ser Leu Ala  
 125 130 135  
 Asn Val Pro Phe Leu Ile Leu Gly Asn Lys Ile Asp Ile Pro Tyr  
 140 145 150  
 Ala Ala Ser Glu Glu Glu Leu Arg Tyr His Leu Gly Leu Ser Asn  
 155 160 165  
 Phe Thr Thr Gly Lys Gly Lys Val Asn Leu Gly Asp Ser Asn Val  
 170 175 180  
 Arg Pro Leu Glu Val Phe Met Cys Ser Val Val Arg Lys Met Gly  
 185 190 195  
 Tyr Gly Asp Gly Phe Lys Trp Val Ser Gln Tyr Ile Lys  
 200 205

<210> 319  
 <211> 144  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:457478.1.orf3:2000SEP08

<400> 319  
 Ala Ser Gly Ser Ala Arg Ala Ser Trp Ser Thr Trp Ser Ser Gly  
 1 5 10 15  
 Ala Ala Gly Pro Pro Asn Ile Thr Ala Gly Ser Arg Arg Arg Thr  
 20 25 30  
 Ser Trp Thr Arg Gly Cys Ser Trp Pro Ser Arg Arg Ser Glu Asp  
 35 40 45  
 Ala Asp Ser Thr Gly Glu Gly Val Gly Glu Gly Arg Trp Leu Val  
 50 55 60  
 Gly Val Val Leu Gly Ala Ala Arg Leu Pro Gly Gly Trp Ser Trp  
 65 70 75  
 Gly Cys Gly Leu Arg Ser Leu Ser Lys Ile Val Ser Arg Ala Ile  
 80 85 90

```

Ser Val Ala Gly Cys Lys Ser Pro Gln Pro Ala Pro Leu Pro Pro
          95                      100                      105
Ser Pro Ser Gln Pro Val Ser Ile Pro Glu Lys Asn Leu Ser Lys
          110                      115                      120
Cys Cys Leu Gly Val Ser Asp Ser His Thr Cys Lys Asn Gln Leu
          125                      130                      135
Phe Gln Leu Leu Gln Ile Lys Met Leu
          140

```

&lt;210&gt; 320

&lt;211&gt; 302

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:358719.1.orf3:2000SEP08

&lt;400&gt; 320

```

Ser Phe Leu Leu Arg Ala Ala Pro Thr Pro Gly Leu Lys Met Arg
  1          5                      10                      15
His Ile Ile Cys His Gly Gly Val Ile Thr Glu Glu Met Ala Ala
          20                      25                      30
Ser Leu Leu Asp Gln Leu Ile Glu Glu Val Leu Ala Asp Asn Leu
          35                      40                      45
Pro Pro Pro Ser His Phe Glu Pro Pro Thr Leu His Glu Leu Tyr
          50                      55                      60
Asp Leu Asp Val Thr Ala Pro Glu Asp Pro Asn Glu Glu Ala Val
          65                      70                      75
Ser Gln Ile Phe Pro Asp Ser Val Met Leu Ala Val Gln Glu Gly
          80                      85                      90
Ile Asp Leu Leu Thr Phe Pro Pro Ala Pro Gly Ser Pro Glu Pro
          95                      100                     105
Pro His Leu Ser Arg Gln Pro Glu Gln Pro Glu Gln Arg Ala Leu
          110                     115                     120
Gly Pro Val Ser Met Pro Asn Leu Val Pro Glu Val Ile Asp Leu
          125                     130                     135
Thr Cys His Glu Ala Gly Phe Pro Pro Ser Asp Asp Glu Asp Glu
          140                     145                     150
Glu Gly Glu Glu Phe Val Leu Asp Tyr Val Glu His Pro Gly His
          155                     160                     165
Gly Cys Arg Ser Cys His Tyr His Arg Arg Asn Thr Gly Asp Pro
          170                     175                     180
Asp Ile Met Cys Ser Leu Cys Tyr Met Arg Thr Cys Gly Met Phe
          185                     190                     195
Val Tyr Ser Pro Val Ser Glu Pro Glu Pro Glu Pro Glu Pro Glu
          200                     205                     210
Pro Glu Pro Ala Arg Pro Thr Arg Arg Pro Lys Met Ala Pro Ala
          215                     220                     225
Ile Leu Arg Arg Pro Thr Ser Pro Val Ser Arg Glu Cys Asn Ser
          230                     235                     240
Ser Thr Asp Ser Cys Asp Ser Gly Pro Ser Asn Thr Pro Pro Glu
          245                     250                     255
Ile His Pro Val Val Pro Leu Cys Pro Ile Lys Pro Val Ala Val
          260                     265                     270
Arg Val Gly Gly Arg Arg Gln Ala Val Glu Cys Ile Glu Asp Leu
          275                     280                     285
Leu Asn Glu Pro Gly Gln Pro Leu Asp Leu Ser Cys Lys Arg Pro
          290                     295                     300
Arg Pro

```

&lt;210&gt; 321

<211> 162  
 <212> PRT  
 <213> Homo sapiens.

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:105160.5.orf1:2000SEP08

<400> 321  
 Ala Thr Glu Arg His His Arg Asp Val Val Glu Leu Leu Ile Lys  
   1                  5                  10                  15  
 Tyr Gly Ala Asp Val His Ala Phe Ser Lys Phe Asp Lys Ser Ala  
                   20                  25                  30  
 Phe Asp Ile Ala Leu Glu Lys Asn Asn Ala Glu Ile Leu Val Ile  
                   35                  40                  45  
 Leu Gln Glu Ala Met Gln Asn Gln Val Asn Val Asn Pro Glu Arg  
                   50                  55                  60  
 Ala Asn Pro Val Thr Asp Pro Val Ser Met Ala Ala Pro Phe Ile  
                   65                  70                  75  
 Phe Thr Ser Gly Glu Val Val Asn Leu Ala Ser Leu Ile Ser Ser  
                   80                  85                  90  
 Thr Asn Thr Lys Thr Thr Ser Ala Asn Thr Glu Glu Ile Ile Glu  
                   95                  100                 105  
 Gly Asn Ser Val Asp Ser Ser Ile Gln Gln Val Met Gly Ser Gly  
                  110                 115                 120  
 Gly Gln Arg Val Ile Thr Ile Val Thr Asp Gly Val Pro Leu Gly  
                  125                 130                 135  
 Asn Ile Gln Thr Ser Ile Pro Thr Gly Gly Ile Gly Gln Pro Phe  
                  140                 145                 150  
 Ile Val Thr Val Gln Asp Gly Gln Gln Gly Arg Gln  
                  155                 160

<210> 322  
 <211> 195  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:400705.1.orf1:2000SEP08

<400> 322  
 Arg Arg Ala Pro Ala Gly Pro Met Ser Phe Ala Thr Leu Arg Pro  
   1                  5                  10                  15  
 Ala Pro Pro Gly Arg Tyr Leu Tyr Pro Glu Val Ser Pro Leu Ser  
                   20                  25                  30  
 Glu Asp Glu Asp Arg Gly Ser Asp Ser Ser Gly Ser Asp Glu Lys  
                   35                  40                  45  
 Pro Cys Arg Val His Ala Ala Arg Cys Gly Leu Gln Gly Ala Arg  
                   50                  55                  60  
 Arg Arg Ala Gly Gly Arg Arg Ala Gly Gly Gly Gly Pro Gly Gly  
                   65                  70                  75  
 Arg Pro Gly Arg Glu Pro Arg Gln Arg His Thr Ala Asn Ala Arg  
                   80                  85                  90  
 Glu Arg Asp Arg Thr Asn Ser Val Asn Thr Ala Phe Thr Ala Leu  
                   95                 100                 105  
 Arg Thr Leu Ile Pro Thr Glu Pro Ala Asp Arg Lys Leu Ser Lys  
                  110                 115                 120  
 Ile Glu Thr Leu Arg Leu Ala Phe Ser Tyr Ile Ser His Leu Gly  
                  125                 130                 135  
 Asn Val Leu Leu Ala Gly Glu Ala Cys Gly Thr Asp Ser Pro Ala  
                  140                 145                 150  
 Thr Pro Gly Pro Pro Ser Ser Thr Arg Arg Ala Pro Ala Ala Pro

				155					160				165
Arg	Arg	Arg	Pro	Arg	Arg	Leu	Pro	Pro	Ala	Thr	Ala	Arg	Thr
				170					175				180
Ser	Pro	Asn	Arg	Ser	Ala	Pro	Ser	Ala	Ser	Ala	Thr	Arg	Glu
				185					190				195

&lt;210&gt; 323

&lt;211&gt; 124

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:221977.1.orf3:2000SEP08

&lt;400&gt; 323

Val	Tyr	Cys	Gly	Leu	Arg	Arg	Ser	Arg	His	Cys	Cys	Arg	Pro	Leu
1				5					10					15
Pro	Ala	Thr	Ile	Phe	Leu	Gly	Ala	Arg	Gly	Pro	Arg	Val	Thr	Gly
				20					25					30
Glu	His	Leu	Thr	Ala	Pro	Pro	Tyr	Pro	Ala	Gln	Met	Pro	Glu	Arg
				35					40					45
Pro	Pro	Trp	Leu	Ser	Cys	Pro	Ser	Trp	Phe	Arg	Thr	Ala	Gly	Glu
				50					55					60
Arg	Pro	Gln	Gln	Leu	Cys	Cys	Pro	Ala	Val	Phe	Gly	Gly	Gly	
				65					70					75
Glu	Gly	Trp	Thr	Cys	Gly	Thr	Leu	Met	Gln	Pro	Gln	Arg	Gln	Gly
				80					85					90
Arg	Gly	Arg	Gly	Gly	Lys	Gly	Trp	Tyr	Val	Glu	Gly	Asn	Gly	Trp
				95					100					105
Trp	Asp	Gln	Gly	Pro	Asn	Ala	Asn	Lys	Asp	Trp	Thr	Val	Leu	Gln
				110					115					120
Lys	Lys	Lys	Lys											

&lt;210&gt; 324

&lt;211&gt; 262

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:898771.1.orf1:2000SEP08

&lt;400&gt; 324

Gly	Arg	Gly	Arg	Arg	Ile	Lys	Ile	Gln	Phe	Thr	Ser	Leu	Tyr
1				5				10					15
His	Lys	Glu	Glu	Ala	Pro	Ala	Ser	Pro	Leu	Arg	Pro	Leu	Tyr
				20				25					30
Gln	Ile	Ser	Pro	Leu	Lys	Ile	His	Ile	Pro	Glu	Pro	Asp	Leu
				35				40					45
Ser	Met	Val	Ser	Pro	Val	Pro	Ser	Pro	Thr	Gly	Thr	Ile	Ser
				50				55					60
Pro	Asn	Ser	Cys	Pro	Ala	Ser	Pro	Arg	Gly	Ala	Gly	Ser	Ser
				65				70					75
Tyr	Arg	Phe	Val	Gln	Asn	Val	Thr	Ser	Asp	Leu	Gln	Leu	Ala
				80				85					90
Glu	Phe	Ala	Ala	Lys	Ala	Ala	Ser	Glu	Gln	Gln	Ala	Asp	Thr
				95				100					105
Gly	Gly	Asp	Ser	Pro	Lys	Asp	Glu	Ser	Lys	Pro	Pro	Phe	Ser
				110				115					120
Ala	Gln	Leu	Ile	Val	Gln	Ala	Ile	Ser	Ser	Ala	Gln	Asp	Arg

	125		130		135
Leu Thr Leu Ser	Gly Ile Tyr Ala His	Ile Thr Lys His Tyr	Pro		
	140		145		150
Tyr Tyr Arg Thr	Ala Asp Lys Gly Trp	Gln Asn Ser Ile Arg	His		
	155		160		165
Asn Leu Ser Leu	Asn Arg Tyr Phe Ile	Lys Val Pro Arg Ser	Gln		
	170		175		180
Glu Glu Pro Gly	Lys Gly Ser Phe Trp	Arg Ile Asp Pro Ala	Ser		
	185		190		195
Glu Ala Lys Leu	Val Glu Gln Ala Phe	Arg Lys Arg Arg Gln	Arg		
	200		205		210
Gly Val Ser Cys	Phe Arg Thr Pro Phe	Gly Pro Leu Ser Ser	Arg		
	215		220		225
Ser Ala Pro Ala	Ser Pro Thr His Pro	Gly Leu Met Ser Pro	Arg		
	230		235		240
Ser Gly Gly Leu	Gln Thr Pro Glu Cys	Leu Ser Arg Glu Gly	Ser		
	245		250		255
Pro Ile Pro His	Asn Ile Arg				
	260				

&lt;210&gt; 325

&lt;211&gt; 98

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:457478.1.orf2:2000SEP08

&lt;400&gt; 325

Gln His Cys Gly	Glu Gly Val Gly	Glu Gly Arg Trp	Leu Val Gly
1	5	10	15
Val Val Leu Gly	Ala Ala Arg Leu	Pro Gly Gly Trp	Ser Trp Gly
	20	25	30
Cys Gly Leu Arg	Ser Leu Ser Lys	Ile Val Ser Arg	Ala Ile Ser
	35	40	45
Val Ala Gly Cys	Lys Ser Pro Gln	Pro Ala Pro Leu	Pro Pro Ser
	50	55	60
Pro Ser Gln Pro	Val Ser Ile Pro	Glu Lys Asn Leu	Ser Lys Cys
	65	70	75
Cys Phe Gly Val	Ser Asp Ser His	Thr Cys Lys Asn	Gln Leu Phe
	80	85	90
Gln Leu Leu Gln	Ile Lys Met Leu		
	95		

&lt;210&gt; 326

&lt;211&gt; 408

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:125140.1.orf3:2000SEP08

&lt;400&gt; 326

Gly Ser Ser Ala	Ser Trp Ala Ser	Tyr Ala Thr Ala	Pro Cys Trp
1	5	10	15
Trp Ala Thr Arg	Ala Ser Arg Pro	Thr Val Pro Cys	Trp Pro Arg
	20	25	30
Ala Ala Ser Thr	Ser Ile Ser Ser	Thr Gly Thr Gly	Pro Arg Ala
	35	40	45
Val Ala Thr Arg	Cys Gly Ser Thr	Ala Thr Ser Ser	Arg Arg Pro
	50	55	60

```

Pro Ser Ala Ala Tyr Trp Thr Ser Cys Thr Arg Ala Ala Trp Thr
65 70 75
Trp Arg Ser Leu Pro Val Glu Asp Val Leu Ala Ala Ala Ser Tyr
80 85 90
Leu His Met Tyr Asp Ile Val Lys Val Cys Lys Gly Arg Leu Gln
95 100 105
Glu Lys Asp Arg Ser Leu Asp Pro Gly Asn Pro Ala Pro Gly Ala
110 115 120
Glu Pro Ala Gln Pro Pro Cys Pro Trp Pro Val Trp Thr Ala Asp
125 130 135
Leu Cys Pro Ala Ala Arg Lys Ala Lys Leu Pro Pro Phe Gly Val
140 145 150
Lys Ala Ala Leu Pro Pro Arg Ala Ser Gly Pro Pro Pro Cys Gln
155 160 165
Val Pro Glu Glu Ser Asp Gln Ala Leu Asp Leu Ser Leu Lys Ser
170 175 180
Gly Pro Arg Gln Glu Arg Val His Pro Pro Cys Val Leu Gln Thr
185 190 195
Pro Leu Cys Ser Gln Arg Gln Pro Gly Ala Gln Pro Leu Val Lys
200 205 210
Asp Glu Arg Asp Ser Leu Ser Lys Gln Glu Glu Ile Ser Ser Ser
215 220 225
Arg Ser Pro His Ser Pro Pro Lys Pro Pro Pro Val Pro Ala Ala
230 235 240
Lys Gly Leu Val Val Gly Leu Gln Pro Leu Pro Leu Ser Gly Glu
245 250 255
Gly Ser Arg Glu Leu Glu Leu Gly Ala Gly Arg Leu Ala Ser Glu
260 265 270
Asp Glu Leu Gly Pro Gly Gly Pro Leu Cys Ile Cys Pro Leu Cys
275 280 285
Ser Lys Leu Phe Pro Ser Ser His Val Leu Gln Leu His Leu Ser
290 295 300
Ala His Phe Arg Glu Arg Asp Ser Thr Arg Ala Arg Ile Ser Pro
305 310 315
Asp Gly Val Ala Pro Thr Cys Pro Leu Cys Gly Lys Thr Phe Ser
320 325 330
Cys Thr Tyr Thr Leu Lys Arg His Glu Arg Thr His Ser Gly Glu
335 340 345
Lys Pro Tyr Thr Cys Val Gln Cys Gly Lys Ser Phe Gln Tyr Ser
350 355 360
His Asn Leu Ser Arg His Thr Val Val His Thr Arg Glu Lys Pro
365 370 375
His Ala Cys Arg Trp Cys Glu Arg Arg Phe Thr Gln Ser Gly Asp
380 385 390
Leu Tyr Arg His Val Arg Lys Phe His Cys Gly Leu Val Lys Ser
395 400 405
Leu Leu Val

```

&lt;210&gt; 327

&lt;211&gt; 59

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:021095.2.orf1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 8

&lt;223&gt; unknown or other



&lt;400&gt; 327

Cys	Leu	Val	Ala	Gly	Thr	Glu	Xaa	Ala	Val	Lys	Met	Ala	Ser	Thr
1				5					10					15
Ser	Arg	Leu	Asp	Ala	Leu	Pro	Arg	Val	Thr	Cys	Pro	Asn	His	Pro
			20						25					30
Asp	Ala	Ile	Leu	Val	Glu	Asp	Tyr	Arg	Ala	Gly	Asp	Met	Ile	Cys
			35						40					45
Pro	Glu	Cys	Gly	Leu	Val	Val	Gly	His	Arg	Ser	Cys	Lys	Phe	
			50						55					

&lt;210&gt; 328

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:888730.1.orf2:2000SEP08

&lt;400&gt; 328

Pro	Arg	Ile	Cys	Ile	Leu	Leu	Pro	Thr	Val	Phe	Gly	Arg	Asn	Arg
1				5					10					15
Pro	Gly	Asp	Ser	His	Phe	Ile	His	Phe	Gln	Pro	Pro	Tyr	Lys	Ser
			20						25					30
Ser	Ser	Lys	Ser	Gln	Pro	Pro	Gly	Pro	Ser	Pro	Cys	Arg	Leu	Arg
			35						40					45
Leu	Leu	Gly	Pro	Ala	Ser	Pro	Val	Phe	Ala	Val	Ser	Ser	Ile	Leu
			50						55					60
Arg	Ile	Ser	His	Asp	Leu	Ser	Asp	Ser	Ser	Glu	Leu	Phe	His	Arg
			65						70					75
Ser	Gly	Ser	Cys	Arg	Glu	Pro	Pro	Gly	Gln	Leu	Ala	Pro	Ala	Gly
			80						85					90
Leu	Leu	His	Leu	Pro	Leu	Ser	Gly	Leu	Leu	Phe	Gly	Ser	Gly	
			95						100					

&lt;210&gt; 329

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:888730.1.orf3:2000SEP08

&lt;400&gt; 329

Asp	Gln	Pro	Ala	Pro	Ser	Ser	Arg	Leu	Ala	Pro	Tyr	Ser	Gly	Ser
1				5					10					15
Ala	Met	Thr	Ser	Gln	Ile	Arg	Gln	Asn	Tyr	Ser	Thr	Glu	Val	Glu
			20						25					30
Ala	Ala	Val	Asn	Arg	Leu	Val	Asn	Leu	His	Leu	Arg	Ala	Ser	Tyr
			35						40					45
Thr	Tyr	Leu	Ser	Leu	Gly	Phe	Phe	Leu	Asp	Arg	Asp	Asp	Val	Ala
			50						55					60
Leu	Glu	Gly	Val	Gly	His	Phe	Leu	Pro	Arg	Ile	Gly	Arg	Gly	Glu
			65						70					75
Ala	Arg	Gly	Arg	Arg	Ala	Ser	Pro	Gln	Val	Ala	Glu	Arg	Thr	Arg
			80						85					90
Gly	Pro	Cys	Thr	Leu	Pro	Gly	Cys	Ala	Glu	Ala	Ile	Ser	Arg	
			95						100					

&lt;210&gt; 330

&lt;211&gt; 168

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:358719.1.orf2:2000SEP08

&lt;400&gt; 330

```

Tyr Asp Leu Arg Arg Asp Gly Pro Arg Arg Ser Gln Arg Arg Ser
 1          5          10          15
Arg Phe Pro Glu Phe Ser Pro Asp Pro Glu Met Leu Ala Val Gln
          20          25          30
Glu Gly Ile Asp Leu Leu Thr Phe Pro Pro Ala Pro Arg Tyr Leu
          35          40          45
Arg Ser Arg Arg His Ala Phe Ala Gly Ser Pro Ser Ser Arg Ser
          50          55          60
Gln Arg Ala Leu Gly Pro Gly Phe Leu Trp Pro Lys Pro Leu Tyr
          65          70          75
Arg Arg Cys Ile Asp Leu Tyr Leu Pro Arg Gly Trp Leu Ser Thr
          80          85          90
His Val Thr Thr Ile Glu Glu Gly Glu Glu Phe Val Tyr Arg
          95          100          105
Leu Cys Gly Ala Pro Arg Ala Arg Leu Gln Val Leu Ser Leu Ser
          110          115          120
Pro Glu Glu Tyr Gly Gly Pro Arg Tyr Tyr Val Phe Ala Leu Leu
          125          130          135
Tyr Glu Asp Leu Trp His Val Ala Ser Thr Val Leu Cys Leu Asn
          140          145          150
Leu Arg Pro Glu Pro Glu Pro Glu Pro Gly Ala Cys Lys Thr Tyr
          155          160          165
Pro Pro Ser

```

&lt;210&gt; 331

&lt;211&gt; 210

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:351342.3.orf3:2000SEP08

&lt;400&gt; 331

```

Leu Ile Leu Gly Ala Val Ser Ser Arg Arg Val Cys Leu Leu Leu
 1          5          10          15
Arg Arg Gly Leu Cys Cys Lys Leu Cys Cys Gln Leu Gln Val Arg
          20          25          30
Gly His Val Leu His Lys Ala Val Thr Gly Gly Ala Gly Thr Ala
          35          40          45
Trp Thr Gly Trp Ala Gly Val Gly Asp Thr Asp Gly Ala Arg Arg
          50          55          60
Gly Gly Asp Gly Ala Asp His Ala Pro Glu Val Arg Leu Arg Asp
          65          70          75
Val Asp Leu Gln Arg Gly Asp Leu Gly Val Gln Trp Pro Gln Arg
          80          85          90
Gly Gly Trp Gly Leu Phe Phe Val Ile Glu Arg Arg Glu Leu Asp
          95          100          105
Leu Asp Gly Arg Ala Gly Lys Pro Glu Gly Thr Leu Leu Gly Gln
          110          115          120
Leu Gln Gly Gly Arg Ala Ala Ser Leu Glu Gly Pro Val His Glu
          125          130          135
Asp Ala Val Leu Ala Glu Ala Ala Gln Val Glu Val Arg Leu Leu
          140          145          150
Glu Ala Glu Leu Gln Val Ala Pro Arg Asp Glu Ala Gly Gln Ala

```

	155		160		165
His Ala Gln Val	His Arg Ala Leu Arg Arg Val Ala Ala Asp Gly				
	170		175		180
Asp Ala Gly Leu Ala	His Glu Glu Leu Glu Leu Ala Ala Leu Gln				
	185		190		195
Pro Arg Gln Arg Arg	Pro Arg Ala Leu Pro Tyr Arg Gly Gly Arg				
	200		205		210

<210> 332  
 <211> 132  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:256099.2.orf3:2000SEP08

<400> 332	
Ser Arg Arg Gly Ala His Leu Ile Pro Ala Leu Ile Pro Arg Ile	
1 5 10 15	
Pro Ala Ser Ser Ser His Thr Arg Gly Gln Gly Ala Gly Gly Pro	
20 25 30	
Gln Leu Arg Arg Arg Arg Val Pro Pro Glu Arg Ser Pro Arg Pro	
35 40 45	
Gly Arg Val Gly Ala Arg Gly Gly Leu Leu Gly Thr Val Ile Ser	
50 55 60	
Ala Pro Ile Pro Pro Ser Arg Thr Leu Gly Pro Val Gln Arg Ala	
65 70 75	
Leu Leu Gln Ser Gly Ala Leu Gly Gly His Arg Ser Ala Thr His	
80 85 90	
Tyr His Pro Thr Tyr Val Glu Leu Ser His Leu Pro Cys Arg Gly	
95 100 105	
Ala Pro Ala Leu Gly Ala Val Arg Leu Leu Asp Lys Pro Ile Asp	
110 115 120	
Asn Gly Thr Gln Arg Ala Cys Gly Ala Arg Ile Leu	
125 130	

<210> 333  
 <211> 85  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:2051991.1.orf1:2000SEP08

<400> 333	
Gly Ser Leu Trp Leu Tyr His Ala His Glu Arg Lys Gly Gln Met	
1 5 10 15	
Lys Ala Gly Arg Ser Thr Ser Gln Ala Gln Trp Glu Gly His Met	
20 25 30	
Gln Gly Ala Ala Trp Trp His Ala Gly Phe Pro Arg Ser Phe Gln	
35 40 45	
Phe Ala Leu Arg Ala Val Arg Ser Trp Arg Pro Ala Leu His Phe	
50 55 60	
Ser Trp Gln Val Ser Thr Glu Ser His Glu Ser Cys Gly Phe His	
65 70 75	
Pro Gln Ala Ser Thr Phe Ala Ala Asn Leu	
80 85	

<210> 334  
 <211> 119

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:980769.1.orf3:2000SEP08

&lt;400&gt; 334

Arg	Ser	Gln	Gly	Lys	Ser	Leu	Ser	Gln	Glu	Val	Trp	Lys	Glu	His
1				5					10					15
Trp	Ser	Phe	Cys	Ile	Arg	Asp	Leu	Gly	Arg	Ala	Ser	Gly	Lys	Cys
				20					25					30
Trp	Thr	Val	Ser	Pro	Ser	Pro	Ala	Pro	Thr	Pro	Leu	Pro	Ser	Thr
				35					40					45
Ser	Ser	Pro	Pro	Ser	Pro	Pro	Asp	Arg	Leu	Ser	Ala	Ala	Ala	Arg
				50					55					60
Ala	Ala	Phe	Arg	Pro	Ala	Ala	Val	Ala	Gln	Ala	Arg	Pro	Leu	Leu
				65					70					75
Leu	Val	Leu	Gly	Ala	Val	Ala	Gln	Gly	Arg	Ala	Val	Leu	His	Gln
				80					85					90
Leu	Leu	Pro	Leu	Gln	Arg	Val	Gln	Leu	Gln	Ile	His	Ser	Arg	Arg
				95					100					105
Pro	Ser	Pro	Ile	Glu	Met	Pro	Gln	Gly	Pro	His	Arg	Gly	Leu	
				110					115					

&lt;210&gt; 335

&lt;211&gt; 122

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:332474.3.orf3:2000SEP08

&lt;400&gt; 335

Ala	Ser	Ser	Ala	Glu	Leu	Leu	Leu	Tyr	Thr	Cys	Phe	Pro	Thr	Thr
1				5					10					15
Ala	Arg	Lys	Asn	His	Gly	Pro	Gly	Val	Thr	Val	Ile	Arg	Gly	Arg
				20					25					30
Gly	Cys	Gly	Leu	His	Pro	Glu	Arg	Val	Ala	Ala	Ala	Gly	Pro	Gly
				35					40					45
Ala	Glu	Asp	Pro	Val	Pro	Gly	Cys	Asp	Ala	Gly	Glu	Leu	Gln	Pro
				50					55					60
Pro	Ala	Leu	Cys	Gly	Leu	Asn	Cys	Leu	Leu	Ala	Trp	Lys	Ala	Gln
				65					70					75
Leu	Ala	Gln	His	Lys	Ser	Phe	Asn	Arg	Phe	Ser	Pro	Arg	Gly	Cys
				80					85					90
Gln	Val	Ser	Lys	Pro	Ala	Val	Ile	Ser	Ser	Leu	Glu	Gln	Gly	Lys
				95					100					105
Glu	Pro	Trp	Met	Glu	Glu	Glu	Glu	Ile	Arg	Thr	Trp	Ser	Phe	Pro
				110					115					120
Glu	Ser													

&lt;210&gt; 336

&lt;211&gt; 137

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1087707.1.orf1:2000SEP08

&lt;400&gt; 336

```

Leu His His Ser Pro Cys Tyr Pro Val Thr Cys Arg Tyr Trp Glu
 1          5          10          15
Ile His Arg Glu Glu Gly Gly Thr Ser Gly Gly Trp Glu Met Arg
          20          25          30
Val Leu Thr Phe Arg Asp Val Ala Val Glu Phe Ser Pro Glu Glu
          35          40          45
Trp Glu Cys Leu Asp Ser Ala Gln Gln Arg Leu Tyr Arg Asp Val
          50          55          60
Met Leu Glu Asn Tyr Gly Asn Leu Phe Ser Leu Gly Leu Ala Ile
          65          70          75
Phe Lys Pro Asp Leu Ile Thr Tyr Leu Glu Gln Arg Lys Glu Pro
          80          85          90
Trp Asn Ala Arg Arg Gln Lys Thr Val Ala Lys His Pro Gly Tyr
          95          100          105
Tyr Asp Val Cys His Glu Asp Tyr Glu Tyr Asn Trp Ser Tyr Met
          110          115          120
Phe Leu Asn Ser Glu Gln Leu Phe Thr Lys Phe Tyr Pro Thr Phe
          125          130          135
Phe Cys

```

&lt;210&gt; 337

&lt;211&gt; 98

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:415349.1.orf2:2000SEP08

&lt;400&gt; 337

```

Lys Asn Ile Leu Val Ser Pro Glu Phe Ser Val Phe His Asn Asn
 1          5          10          15
Gly Asn Ile Thr Glu Glu Ser Met Ala Thr Leu Leu Ile Gln Gly
          20          25          30
Val Gln Gln Thr Ala Arg Val Leu Asn Val Cys Ser Arg Asn Gln
          35          40          45
Phe Cys Thr Leu His Ala Cys Phe Pro Ala Pro Pro Leu Leu Leu
          50          55          60
Leu Arg Thr Arg Gly Arg Leu Gly Val Pro Gln Gln Glu Glu Ala
          65          70          75
Ala Leu Pro Ser Asp Ser Ala Pro Ser Leu Asn Val Ala Pro Val
          80          85          90
Trp Val Asp Val Asp Ala Arg Pro
          95

```

&lt;210&gt; 338

&lt;211&gt; 166

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:132420.2.orf2:2000SEP08

&lt;400&gt; 338

```

Arg Trp Val Ser Val Ser His Thr Arg Val Lys Leu Phe Leu Glu
 1          5          10          15
Gln Met Ala Met Arg Glu Leu Asn Ala Asp Ser Cys Ser Ser Pro
          20          25          30
Gln Met Gly Ala Met Trp Glu Thr Ser Gly Ser Val Lys Glu Asn
          35          40          45

```

```

Ser Ser Gln Ser Lys Lys Tyr Ser Thr Lys Ile Glu Asn Leu Gly
      50      55      60
Pro Glu Ser Ala Cys Arg His Phe Trp Ser Phe Arg Tyr His Glu
      65      70      75
Ala Thr Gly Pro Leu Glu Thr Ile Ser Gln Leu Gln Lys Leu Cys
      80      85      90
His Gln Trp Leu Arg Pro Glu Ile His Ser Lys Glu Gln Ile Leu
      95     100     105
Glu Met Leu Val Leu Glu Gln Phe Leu Ser Ile Leu Pro Lys Glu
     110     115     120
Thr Gln Asn Trp Val Gln Lys His His Pro Gln Asn Val Lys Gln
     125     130     135
Ala Leu Val Leu Val Glu Phe Leu Gln Arg Glu Pro Asp Gly Thr
     140     145     150
Lys Asn Glu Val Arg Arg Lys Ile Ala Tyr Met Tyr Asn Glu Val
     155     160     165
Gly

```

```

<210> 339
<211> 65
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LG:394201.1.orf2:2000SEP08

```

```

<400> 339
Phe Leu Ser Thr Val Asp Phe Leu Val Thr Glu Glu Ala Gln Ala
  1      5      10      15
Leu Ala Lys Ala Phe Pro Ser Leu Thr Thr Leu Leu Trp Leu Leu
     20      25      30
Phe Ser Met Asn Ser Asp Val Val Ser Asp Met Asn Ser Glu Arg
     35      40      45
Ile Ser His Asn His Tyr Thr Gly Met Val Ser Phe Pro Cys Glu
     50      55      60
Phe Ser Gly Val Lys
     65

```

```

<210> 340
<211> 198
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LG:1060884.1.orf3:2000SEP08

```

```

<400> 340
Val Gly Phe Gln Phe Pro Lys Pro Glu Met Ile Cys Gln Leu Glu
  1      5      10      15
Asn Trp Asp Glu Gln Trp Ile Leu Asp Leu Pro Arg Thr Gly Asn
     20      25      30
Arg Lys Ala Ser Gly Ser Ala Cys Pro Gly Ser Glu Ala Arg His
     35      40      45
Lys Met Lys Lys Leu Thr Pro Lys Gln Lys Phe Ser Glu Asp Leu
     50      55      60
Glu Ser Tyr Lys Ile Ser Val Val Met Gln Glu Ser Ala Glu Lys
     65      70      75
Leu Ser Glu Lys Leu His Lys Cys Lys Glu Phe Val Asp Ser Cys
     80      85      90
Arg Leu Thr Phe Pro Thr Ser Gly Asp Glu Tyr Ser Arg Gly Phe

```

	95		100		105
Leu Gln Asn Leu Asn Leu Ile Gln Asp		Gln Asn Ala Gln Thr Arg			
	110		115		120
Trp Lys Gln Gly Arg Tyr Asp Glu Asp		Gly Lys Pro Phe Asn Gln			
	125		130		135
Arg Ser Leu Leu Leu Gly His Glu Arg		Ile Leu Thr Arg Ala Lys			
	140		145		150
Ser Tyr Glu Cys Ser Glu Cys Gly Lys		Val Ile Arg Arg Lys Ala			
	155		160		165
Trp Phe Asp Gln His Gln Arg Ile His		Phe Leu Glu Asn Pro Phe			
	170		175		180
Glu Cys Lys Val Cys Gly Gln Ala Phe		Arg Gln Arg Ser Ala Leu			
	185		190		195
Thr Val Pro					

&lt;210&gt; 341

&lt;211&gt; 439

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:242191.1.orf2:2000SEP08

&lt;400&gt; 341

Ser Gly Glu Gln Ser Gln Thr Ala Gly	Leu Gln Glu Val Glu Ile
1	5 10 15
Leu Trp Leu Asn Arg Arg Cys Arg Thr	Trp Ser Ser Ile Asp Asn
	20 25 30
Ser Cys Phe Asp Phe Pro Gly Glu Glu	Trp Met Ile Arg Lys Val
	35 40 45
Lys Val Glu Asp Glu Asp Gln Glu Ala	Glu Glu Glu Val Glu Trp
	50 55 60
Pro Gln His Leu Ser Leu Leu Pro Ser	Pro Phe Pro Ala Pro Asp
	65 70 75
Leu Gly His Leu Ala Ala Ala Tyr Lys	Leu Glu Pro Gly Ala Pro
	80 85 90
Gly Ala Leu Ser Gly Leu Ala Leu Ser	Gly Trp Gly Pro Met Pro
	95 100 105
Glu Lys Pro Tyr Gly Cys Gly Glu Cys	Glu Arg Arg Phe Arg Asp
	110 115 120
Gln Leu Thr Leu Arg Leu His Gln Arg	Leu His Arg Gly Glu Gly
	125 130 135
Pro Cys Ala Cys Pro Asp Cys Gly Arg	Ser Phe Thr Gln Arg Ala
	140 145 150
His Met Leu Leu His Gln Arg Ser His	Arg Gly Glu Arg Pro Phe
	155 160 165
Pro Cys Ser Glu Cys Asp Lys Arg Phe	Ser Lys Lys Ala His Leu
	170 175 180
Thr Arg His Leu Arg Thr His Thr Gly	Glu Arg Pro Tyr Pro Cys
	185 190 195
Ala Glu Cys Gly Lys Arg Phe Ser Gln	Lys Ile His Leu Gly Ser
	200 205 210
His Gln Lys Thr His Thr Gly Glu Arg	Pro Phe Pro Cys Thr Glu
	215 220 225
Cys Glu Lys Arg Phe Arg Lys Lys Thr	His Leu Ile Arg His Gln
	230 235 240
Arg Ile His Thr Gly Glu Arg Pro Tyr	Gln Cys Ala Gln Cys Ala
	245 250 255
Arg Ser Phe Thr His Lys Gln His Leu	Val Arg His Gln Arg Val
	260 265 270
His Gln Thr Ala Gly Pro Ala Arg Pro	Ser Pro Asp Ser Ser Ala

	275		280		285
Ser Pro His Ser	Thr Ala Pro Ser Pro	Thr Pro Ser Phe Pro	Gly		
	290		295		300
Pro Lys Pro Phe	Ala Cys Ser Asp Cys	Gly Leu Ser Phe Gly	Trp		
	305		310		315
Lys Lys Asn Leu	Ala Thr His Gln Cys	Leu His Arg Ser Glu	Gly		
	320		325		330
Arg Pro Phe Gly	Cys Asp Glu Cys Ala	Leu Gly Ala Thr Val	Asp		
	335		340		345
Ala Pro Ala Ala	Lys Pro Leu Ala Ser	Ala Pro Gly Gly Pro	Gly		
	350		355		360
Cys Gly Pro Gly	Ser Asp Pro Val Val	Pro Gln Arg Ala Pro	Ser		
	365		370		375
Gly Glu Arg Ser	Phe Phe Cys Pro Asp	Cys Gly Arg Gly Phe	Ser		
	380		385		390
His Gly Gln His	Leu Ala Arg His Pro	Arg Val His Thr Gly	Glu		
	395		400		405
Arg Pro Phe Ala	Cys Thr Gln Cys Asp	Arg Arg Phe Gly Ser	Arg		
	410		415		420
Pro Asn Leu Val	Ala His Ser Arg Ala	His Ser Gly Ala Arg	Pro		
	425		430		435
Phe Ala Cys Ala					

&lt;210&gt; 342

&lt;211&gt; 404

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1063762.3.orf1:2000SEP08

&lt;400&gt; 342

Gly Pro Phe Gly	Gln Pro Leu Arg Arg	Ser Glu Leu Glu Thr	Thr
1	5	10	15
Glu Pro Ser Ala	Gly Ala Gly Gly Ala	Ala Arg Gly Ser Gly	Trp
	20	25	30
Ala Gly Gly Arg	Glu Ile Arg Ser Pro	Pro Glu Ala Val Ile	Leu
	35	40	45
Thr Pro Ala Pro	Glu Pro Gly Pro Gly	Ala Gly Met Ala Pro	
	50	55	60
Arg Pro Pro Thr	Ala Ala Pro Gln Glu	Ser Val Thr Phe Lys	Asp
	65	70	75
Val Ser Val Asp	Phe Thr Gln Glu Glu	Trp Tyr His Val Asp	Pro
	80	85	90
Ala Gln Arg Ser	Leu Tyr Arg Asp Val	Met Leu Glu Asn Tyr	Ser
	95	100	105
His Leu Val Ser	Leu Gly Tyr Gln Val	Ser Lys Pro Glu Val	Ile
	110	115	120
Phe Lys Leu Glu	Gln Gly Glu Glu Pro	Trp Ile Ser Glu Gly	Glu
	125	130	135
Ile Gln Arg Pro	Phe Tyr Pro Asp Trp	Lys Thr Arg Pro Glu	Val
	140	145	150
Lys Ser Ser His	Leu Gln Gln Asp Val	Ser Glu Val Ser His	Cys
	155	160	165
Thr His Asp Leu	Leu His Ala Thr Leu	Glu Asp Ser Trp Asp	Val
	170	175	180
Ser Ser Gln Leu	Asp Arg Gln Gln Glu	Asn Trp Lys Arg His	Leu
	185	190	195
Gly Ser Glu Ala	Ser Thr Gln Lys Lys	Ile Ile Thr Pro Gln	Glu
	200	205	210
Asn Phe Glu Gln	Asn Lys Phe Gly Glu	Asn Ser Arg Leu Asn	Thr



	215		220		225
Asn Leu Val Thr	Gln Leu Asn Ile Pro	Ala Arg Ile Arg Pro	Ser		
	230		235		240
Glu Cys Glu Thr	Leu Gly Ser Asn Leu	Gly His Asn Ala Asp	Leu		
	245		250		255
Leu Asn Glu Asn	Asn Ile Leu Ala Lys	Lys Lys Pro Tyr Lys	Cys		
	260		265		270
Asp Lys Cys Arg	Lys Ala Phe Ile His	Arg Ser Ser Leu Thr	Lys		
	275		280		285
His Glu Lys Thr	His Lys Gly Glu Gly	Ala Phe Pro Asn Gly	Thr		
	290		295		300
Asp Gln Gly Ile	Tyr Pro Gly Lys Lys	His His Glu Cys Thr	Asp		
	305		310		315
Cys Gly Lys Thr	Phe Leu Trp Lys Thr	Gln Leu Thr Glu His	Gln		
	320		325		330
Arg Ile His Thr	Gly Glu Lys Pro Phe	Glu Cys Asn Val Cys	Gly		
	335		340		345
Lys Ala Phe Arg	His Ser Ser Ser Leu	Gly Gln His Glu Asn	Ala		
	350		355		360
His Thr Gly Glu	Lys Pro Tyr Gln Cys	Ser Leu Cys Gly Lys	Ala		
	365		370		375
Phe Gln Arg Ser	Ser Ser Leu Val Gln	His Gln Arg Ile His	Thr		
	380		385		390
Gly Glu Lys Pro	Tyr Arg Cys Lys Ser	Met Trp Glu Val Leu			
	395		400		

&lt;210&gt; 343

&lt;211&gt; 139

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1100856.1.orf2:2000SEP08

&lt;400&gt; 343

Ser Leu Ser Leu Ile	Met Gly Arg Ser	Arg Arg Thr Gly	Ala His
1	5	10	15
Arg Ala His Ser	Leu Ala Arg Gln Met	Lys Ala Lys Lys Arg	Arg
	20	25	30
Pro Asp Leu Asp	Glu Ile His Arg Glu	Leu Arg Pro Gln	Ala Leu
	35	40	45
Pro Arg Pro Lys	Arg Glu Arg Asp Ala	Glu Pro Asp Pro	Asp Leu
	50	55	60
Pro Gly Gly Gly	Leu His Arg Cys Leu	Ala Cys Ala Arg	Tyr Phe
	65	70	75
Ile Asp Ser Ala	Asn Leu Lys Thr His	Phe Arg Ser Lys	Asp His
	80	85	90
Lys Lys Arg Leu	Lys Gln Leu Ser Val	Glu Pro Tyr Ser	Gln Glu
	95	100	105
Glu Ala Glu Arg	Ala Ala Gly Met Gly	Ser Tyr Val Gln	Pro Gln
	110	115	120
Arg Leu Gly Val	Pro Thr Glu Val Ser	Thr Glu Ile Pro	Glu Met
	125	130	135
Asp Thr Ser Thr			

&lt;210&gt; 344

&lt;211&gt; 169

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

<221> misc\_feature  
 <223> Incyte ID No: LG:979390.2.orf1:2000SEP08

<220>  
 <221> unsure  
 <222> 23  
 <223> unknown or other

<400> 344  
 Arg Thr Arg Ile Leu Gly Ser Ser Arg Asp Met Val Gly Glu Ala  
 1 5 10 15  
 Glu Gly Leu Ser Leu Ser Ile Xaa Gln His Gly Ala Arg Lys Ser  
 20 25 30  
 Gly Pro Trp Ala Trp Lys Ala Arg Arg Ser Gly Pro Ala Trp Gly  
 35 40 45  
 Thr Lys Ser Cys Ala Thr Ile Pro Ser Thr Cys Gly Gln Leu Pro  
 50 55 60  
 Thr Cys Pro Ala Thr Gly Phe His His Leu Arg Leu Leu Ala Gly  
 65 70 75  
 Ala Ser Leu Ser Leu Cys Pro Val Leu Glu Thr Val Met Cys Met  
 80 85 90  
 His Thr Ser Ser Ser Pro Thr Ala Arg Thr Ala Gly Ile Thr Leu  
 95 100 105  
 Pro Leu Gly Ile Gln Gly Phe Trp Cys Leu Gln Ile Phe Gln Ala  
 110 115 120  
 Pro Thr Met Phe Leu Ser Ser Lys Arg Arg His Pro Arg Trp Lys  
 125 130 135  
 Pro Gly Ser Trp His Ala Arg Pro Ser Cys Ser Pro Ser Ala Asp  
 140 145 150  
 Pro Met Ala Ser Val Gly Ser Ser Pro Lys His Lys Pro Arg Thr  
 155 160 165  
 Ala Arg Pro Gly

<210> 345  
 <211> 128  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1400447.1.orf1:2000SEP08

<400> 345  
 Glu Leu Leu Thr Gln Phe Ser Pro Asn Ser Arg Cys Val Asp Pro  
 1 5 10 15  
 Ser Ser Leu Arg Ser Pro Ala Gly Val Lys Leu His Thr Cys Leu  
 20 25 30  
 His Ser Thr Leu Pro Val Pro Lys Arg Ala His Arg Arg Arg Lys  
 35 40 45  
 Asn Gly Cys Gln Pro Pro Ala Asn His Gly Pro Gly Ile Gly Asp  
 50 55 60  
 Leu Gln Gly Cys Gly Tyr Thr Val His Pro Gly Arg Val Gly Ala  
 65 70 75  
 Ala Glu Pro Arg Pro Glu Gly Pro Val Gln Gly Arg Asp Ala Gly  
 80 85 90  
 Glu Leu Gln Gln Pro Gly Leu Thr Gly Thr Leu Arg Thr Gln Thr  
 95 100 105  
 Arg Tyr Val Phe Pro Ala Arg Lys Lys Gly Ser Val Asp Ala Arg  
 110 115 120  
 Gly His Pro Trp Arg Leu Leu Ser  
 125

<210> 346  
 <211> 146  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1400562.1.orf3:2000SEP08

<400> 346  
 Pro Ala Gly Ile Gly Arg Ser Thr Thr Lys Ser Pro Gly Pro Pro  
 1 5 10 15  
 Gly Ser Leu Glu Met Gly Ser Leu Thr Phe Arg Asp Val Ala Ile  
 20 25 30  
 Glu Phe Ser Leu Glu Trp Gln Cys Leu Asp Thr Ala Gln Gln  
 35 40 45  
 Asn Leu Tyr Arg Asn Val Met Leu Glu Asn Tyr Arg Asn Leu Val  
 50 55 60  
 Phe Leu Gly Ile Ala Ala Phe Lys Pro Asp Leu Ile Ile Phe Leu  
 65 70 75  
 Glu Glu Gly Lys Glu Ser Trp Asn Met Lys Arg His Glu Met Val  
 80 85 90  
 Glu Glu Ser Pro Val Ile Cys Ser His Phe Ala Gln Asp Leu Trp  
 95 100 105  
 Pro Glu Gln Gly Ile Glu Asp Ser Phe Gln Lys Val Ile Leu Arg  
 110 115 120  
 Arg Tyr Lys Ile His His His Ala Cys Glu Leu Gly Pro Ile Met  
 125 130 135  
 Asn His Tyr Pro Thr Cys Gly Gln Met His Ile  
 140 145

<210> 347  
 <211> 250  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1076130.1.orf3:2000SEP08

<220>  
 <221> unsure  
 <222> 12  
 <223> unknown or other

<400> 347  
 Gln Val Lys Ala Asn Glu Ser Asp Cys Arg Ser Xaa Arg Gln Trp  
 1 5 10 15  
 Ala Lys Thr Ser Gly Glu Lys Arg Gly Lys Leu Thr Leu Pro Glu  
 20 25 30  
 Lys Ser Leu Ser Glu Val Leu Ser Gln Gln Arg Pro Cys Leu Gly  
 35 40 45  
 Glu Arg Pro Tyr Lys Tyr Leu Lys Tyr Ser Lys Ser Phe Gly Pro  
 50 55 60  
 Asn Ser Leu Leu Met His Gln Val Ser His Gln Val Glu Asn Pro  
 65 70 75  
 Tyr Lys Cys Ala Asp Cys Gly Lys Ser Phe Ser Arg Ser Ala Arg  
 80 85 90  
 Leu Ile Arg His Arg Arg Ile His Thr Gly Glu Lys Pro Tyr Lys  
 95 100 105  
 Cys Leu Asp Cys Gly Lys Ser Phe Arg Asp Ser Ser Asn Phe Ile  
 110 115 120  
 Thr His Arg Arg Ile His Thr Gly Glu Lys Pro Tyr Gln Cys Gly

	125		130		135
Glu Cys Gly Lys Cys Phe Asn Gln Ser Ser Ser Leu Ile Ile His					
	140		145		150
Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Gln Cys Glu Glu Cys					
	155		160		165
Gly Lys Ser Phe Asn Asn Ser Ser His Phe Ser Ala His Arg Arg					
	170		175		180
Ile His Thr Gly Glu Arg Pro His Val Cys Pro Asp Cys Gly Lys					
	185		190		195
Ser Phe Ser Lys Ser Ser Asp Leu Arg Ala His His Arg Thr His					
	200		205		210
Thr Gly Glu Lys Pro Tyr Gly Cys His Asp Cys Gly Lys Cys Phe					
	215		220		225
Ser Lys Ser Ser Ala Leu Asn Lys His Gly Glu Ile His Ala Arg					
	230		235		240
Glu Lys Leu Leu Thr Gln Ser Ala Pro Lys					
	245		250		

&lt;210&gt; 348

&lt;211&gt; 365

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1064459.1.orf2:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 30-31

&lt;223&gt; unknown or other

&lt;400&gt; 348

Gln Arg Leu Ser Arg Asn Leu Leu Phe Gln Glu Ser Val Thr Phe					
1	5		10		15
Glu Asp Val Ala Val Tyr Phe Thr Gln Asn Gln Trp Ala Ser Xaa					
	20		25		30
Xaa Pro Ala Gln Arg Ala Leu Tyr Gly Glu Val Met Leu Glu Asn					
	35		40		45
Tyr Ala Asn Val Ala Ser Leu Val Ala Phe Pro Phe Pro Lys Pro					
	50		55		60
Ala Leu Ile Ser His Leu Glu Arg Gly Glu Ala Pro Trp Gly Pro					
	65		70		75
Asp Pro Trp Asp Thr Glu Ile Leu Arg Gly Ile Ser Gln Gly Gly					
	80		85		90
Glu Ser Trp Ile Lys Asn Glu Gly Leu Val Ile Lys Gln Glu Ala					
	95		100		105
Ser Glu Glu Thr Glu Leu His Arg Met Pro Val Gly Gly Leu Leu					
	110		115		120
Arg Asn Val Ser Gln His Phe Asp Phe Lys Arg Lys Ala Leu Lys					
	125		130		135
Gln Thr Phe Asn Leu Asn Pro Asn Leu Ile Leu Arg Gly Gly Met					
	140		145		150
Lys Phe Tyr Glu Cys Lys Glu Cys Gly Lys Ile Phe Arg Tyr Asn					
	155		160		165
Ser Lys Leu Ile Arg His Gln Met Ser His Thr Gly Glu Lys Pro					
	170		175		180
Phe Lys Cys Lys Glu Cys Gly Lys Ala Phe Lys Ser Ser Tyr Asp					
	185		190		195
Cys Ile Val His Glu Lys Asn His Ile Gly Glu Gly Pro Tyr Glu					
	200		205		210
Cys Lys Glu Cys Gly Lys Gly Leu Ser Ser Asn Thr Ala Leu Thr					
	215		220		225

Gln His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Lys		
	230	235 240
Glu Cys Gly Lys Ala Phe Arg Arg Ser Ala Ala Tyr Leu Gln His		
	245	250 255
Gln Arg Leu His Thr Gly Glu Lys Leu Tyr Lys Cys Lys Glu Cys		
	260	265 270
Trp Lys Ala Phe Gly Cys Arg Ser Leu Phe Ile Val His Gln Arg		
	275	280 285
Ile His Thr Gly Glu Lys Pro Tyr Gln Cys Lys Glu Cys Gly Lys		
	290	295 300
Ala Phe Thr Gln Lys Ile Ala Ser Ile Gln His Gln Arg Val His		
	305	310 315
Thr Gly Glu Lys Pro Tyr Glu Cys Lys Val Cys Gly Lys Ala Phe		
	320	325 330
Lys Trp Tyr Gly Ser Phe Val Gln His Gln Lys Leu His Pro Val		
	335	340 345
Glu Lys Lys Pro Val Lys Val Leu Gly Pro Ser Leu Val Ser Pro		
	350	355 360
Gln Cys Ser Ser Pro		
	365	

&lt;210&gt; 349

&lt;211&gt; 38

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1079415.14.orf2:2000SEP08

&lt;400&gt; 349

Arg Pro Gly Ile Leu Pro Ser Lys Ser Tyr His Val Tyr Ala Ile		
1	5	10 15
Gly Cys Phe Ser Pro Arg Asn Val Leu His Ser Asn Cys Ile Phe		
	20	25 30
Trp Phe Lys Leu Ser Phe Pro Val		
	35	

&lt;210&gt; 350

&lt;211&gt; 185

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1329431.3.orf2:2000SEP08

&lt;400&gt; 350

Gly Arg Ala Gly Ser Phe Ser Ala Gly Val Gly Val Leu Glu Leu		
1	5	10 15
Arg Ile Pro Gly Leu Glu Ser Arg Arg Gly Ser Ser Trp Leu Pro		
	20	25 30
Phe Ser Ser Gln Ile Cys Leu Leu Glu Thr Ala Pro Ser Ser Arg		
	35	40 45
Glu Ser Gln Lys Glu Asp Met Ala Ala Gly Gln Arg Glu Ala Arg		
	50	55 60
Pro Gln Val Ser Leu Thr Phe Glu Asp Val Ala Val Leu Phe Thr		
	65	70 75
Trp Asp Glu Trp Arg Lys Leu Ala Pro Ser Gln Arg Asn Leu Tyr		
	80	85 90
Arg Asp Val Met Leu Glu Asn Tyr Arg Asn Leu Val Ser Leu Gly		
	95	100 105
Leu Ser Phe Thr Lys Pro Lys Val Ile Ser Leu Leu Gln Gln Gly		

	110		115		120
Glu Asp Pro Trp	Glu Val Glu Lys Asp	Ser Ser Gly Val Ser Ser			
	125		130		135
Leu Gly Cys Lys	Ser Thr Pro Lys Met	Thr Lys Ser Thr Gln Thr			
	140		145		150
Gln Asp Ser Phe	Gln Glu Gln Ile Arg	Asn Arg Leu Pro Arg Asp			
	155		160		165
Glu Pro Trp Asn	Phe Ile Ser Glu Arg	Ser Cys Ile Tyr Glu Glu			
	170		175		180
Lys Leu Lys Lys	Gln				
	185				

&lt;210&gt; 351

&lt;211&gt; 89

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1088431.2.orf1:2000SEP08

&lt;400&gt; 351

Leu Thr Tyr Leu Arg	Lys Lys Leu Arg Gly Arg Gly Lys Lys Glu	
1	5	10 15
Glu Glu Gly Met Ala	Leu Ser Gln Gly Leu Phe Thr Phe Lys Asp	
	20	25 30
Val Ala Ile Glu Phe	Ser Gln Glu Glu Trp Glu Cys Leu Asp Pro	
	35	40 45
Ala Gln Arg Ala Leu	Tyr Arg Asp Val Met Leu Glu Asn Tyr Arg	
	50	55 60
Asn Leu Leu Ser Leu	Asp Glu Asp Asn Ile Pro Pro Glu Asp Gly	
	65	70 75
Ser His Leu Ala Ala	Cys Gly Gln Ser Thr Leu Pro Leu Pro	
	80	85

&lt;210&gt; 352

&lt;211&gt; 123

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1329462.2.orf1:2000SEP08

&lt;400&gt; 352

Val Arg Gly Val Ala	Pro Gly Glu Gly Cys Pro Glu Ala Ser Ala	
1	5	10 15
Gly Pro Arg Gly Asn	Gly Pro Ser Ala Gln Lys Met Leu Leu Leu	
	20	25 30
Ser Leu Phe Phe Pro	Leu Arg Ile Ser Leu Ser Pro Ser Asn His	
	35	40 45
Leu Trp Ser Ala Ser	Pro Arg Cys His Cys Asp Ala Glu Ala Ser	
	50	55 60
Glu Val Ala Gly Ser	Thr Arg Gly Ala Gly Arg Asp Pro Ala Gly	
	65	70 75
Ala Asp Ser Thr Pro	Ser Leu Cys Ser Ala Ser Pro Arg Leu Leu	
	80	85 90
Cys Pro Arg Arg Ser	Gln Leu His Gly Leu Met Pro Ser Gly Ser	
	95	100 105
Ile Leu Lys Ser Ser	Trp Leu Trp Ile Gly Gly Phe Leu Val Ser	
	110	115 120
Thr Asp Leu		

<210> 353  
 <211> 322  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:393468.1.orf2:2000SEP08

<400> 353  
 Tyr Asp Ser Gln Met Ile Asp Leu Cys Asn Val Gly Phe Gln Phe  
 1 5 10 15  
 Tyr Arg Ser Leu Glu His Phe Gly Gly Lys Pro Val Lys Gln Glu  
 20 25 30  
 Pro Ile Lys Pro Ser Ala Val Trp Pro Gln Pro Thr Pro Thr Pro  
 35 40 45  
 Phe Leu Pro Thr Pro Tyr Pro Tyr Tyr Pro Lys Val His Pro Gly  
 50 55 60  
 Leu Met Phe Pro Phe Phe Val Pro Ser Ser Ser Pro Phe Pro Phe  
 65 70 75  
 Ser Arg His Thr Phe Leu Pro Lys Gln Pro Pro Glu Pro Leu Leu  
 80 85 90  
 Pro Arg Lys Ala Glu Pro Gln Glu Ser Glu Glu Thr Lys Gln Lys  
 95 100 105  
 Val Glu Arg Val Asp Val Asn Val Gln Ile Asp Asp Ser Tyr Tyr  
 110 115 120  
 Val Asp Val Gly Gly Ser Gln Lys Arg Trp Gln Phe Pro Thr Cys  
 125 130 135  
 Glu Lys Ser Tyr Thr Ser Lys Tyr Asn Leu Val Thr His Ile Leu  
 140 145 150  
 Gly His Ser Gly Ile Lys Pro His Ala Cys Thr His Cys Gly Lys  
 155 160 165  
 Leu Phe Lys Gln Leu Ser His Leu His Thr His Met Leu Thr His  
 170 175 180  
 Gln Gly Thr Arg Pro His Lys Cys Gln Val Cys His Lys Ala Phe  
 185 190 195  
 Thr Gln Thr Ser His Leu Lys Arg His Met Met Gln His Ser Glu  
 200 205 210  
 Val Lys Pro His Asn Cys Arg Val Cys Gly Arg Gly Leu Ala Tyr  
 215 220 225  
 Pro Ser Glu Leu Lys Ala His Glu Ala Lys His Ala Ser Gly Arg  
 230 235 240  
 Glu Asn Ile Cys Val Glu Cys Gly Leu Asp Phe Pro Thr Leu Ala  
 245 250 255  
 Gln Leu Lys Arg His Leu Thr Thr His Arg Gly Pro Ile Gln Tyr  
 260 265 270  
 Asn Cys Ser Glu Cys Asp Lys Thr Phe Gln Tyr Pro Ser Gln Leu  
 275 280 285  
 Gln Asn His Met Met Lys His Lys Asp Ile Arg Pro Tyr Ile Cys  
 290 295 300  
 Ser Glu Cys Gly Met Glu Phe Val Gln Pro Asp Ser Ser Ser Ser  
 305 310 315  
 Thr Pro Ser Thr Thr Arg Val  
 320

<210> 354  
 <211> 115  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:722577.1.orf1:2000SEP08

&lt;400&gt; 354

```

Val Arg Arg Arg Gly Cys Ser Cys Ser Ser Arg Arg Arg Ile Ser
 1          5          10          15
Ala Pro Pro Ala Arg Ser Leu Val Ser Ser Pro Val Arg Arg Gly
          20          25          30
Ala Thr Met Ala Gly Lys Gly Gly Pro Thr Asn Leu Glu Lys Glu
          35          40          45
Gln Met Phe Gly Met Ala Glu Lys Glu Met Glu Tyr Arg Val Asp
          50          55          60
Leu Phe Asn Arg Leu Thr Lys Thr Cys Phe Glu Lys Cys Val Glu
          65          70          75
Lys Arg Tyr Lys Glu Ala Glu Leu Asn Met Gly Glu Asn Ser Cys
          80          85          90
Ile Asp Arg Cys Val Ser Lys Tyr Trp Gln Val Thr Asn Leu Val
          95          100          105
Gly Gln Met Leu Gly Asn Gln Pro Gln Met
          110          115

```

&lt;210&gt; 355

&lt;211&gt; 108

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:322783.16.orf1:2000SEP08

&lt;400&gt; 355

```

Thr Leu Ile Cys Arg Met Ala Gly Cys Gly Glu Ile Asp His Ser
 1          5          10          15
Ile Asn Met Leu Pro Thr Asn Arg Lys Ala Asn Glu Ser Cys Ser
          20          25          30
Asn Thr Ala Pro Ser Leu Thr Val Pro Glu Cys Ala Ile Cys Leu
          35          40          45
Gln Thr Cys Val His Pro Val Ser Leu Pro Cys Lys His Val Phe
          50          55          60
Cys Tyr Leu Cys Val Lys Gly Ala Ser Trp Leu Gly Lys Arg Cys
          65          70          75
Ala Leu Cys Arg Gln Glu Ile Pro Glu Asp Phe Leu Asp Lys Pro
          80          85          90
Thr Leu Leu Ser Pro Glu Glu Leu Lys Ala Ala Ser Arg Gly His
          95          100          105
Gly Glu Tyr

```

&lt;210&gt; 356

&lt;211&gt; 166

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:901355.2.orf2:2000SEP08

&lt;400&gt; 356

```

Pro Ala Gly Thr Gly Arg Ser Val Ala Gln Thr Pro Gly His Pro
 1          5          10          15
Gly Ser Trp Glu Met Val Asn Val Pro Gly Arg Gly Ser Pro Glu
          20          25          30
Glu Gly Glu Gly Arg Ser Glu Ala Ala Gly Thr Leu Phe Glu Val
          35          40          45
Thr Glu His Leu Thr Gly Trp Ala Pro Ala Pro Val Pro Ser Val
          50          55          60

```



Glu	Gly	Leu	Leu	Thr	Phe	Arg	Asp	Val	Ala	Ile	Glu	Phe	Ser	Arg	
				65					70					75	
Glu	Glu	Trp	Glu	His	Leu	Asp	Ser	Asp	Gln	Lys	Leu	Leu	Tyr	Gly	
				80					85					90	
Asp	Val	Met	Leu	Glu	Asn	Tyr	Gly	Asn	Leu	Val	Ser	Leu	Gly	Leu	
				95					100					105	
Ala	Val	Ser	Lys	Pro	Asp	Leu	Ile	Thr	Phe	Leu	Glu	Gln	Arg	Lys	
				110					115					120	
Glu	Pro	Trp	Asn	Val	Lys	Ser	Ala	Glu	Thr	Val	Ala	Ile	Gln	Pro	
				125					130					135	
Asp	Ile	Phe	Ser	His	Asp	Thr	Gln	Gly	Leu	Leu	Arg	Lys	Lys	Leu	
				140					145					150	
Ile	Glu	Ala	Ser	Phe	Gln	Lys	Val	Ile	Leu	Asp	Gly	Tyr	Gly	Glu	
				155					160					165	

Leu

&lt;210&gt; 357

&lt;211&gt; 114

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:038859.2.orf1:2000SEP08

&lt;400&gt; 357

Ala	Arg	Arg	Ser	Gly	Arg	Ser	Ile	Gly	Ala	Ser	Ala	Leu	Glu	Thr	
1				5					10					15	
Trp	Ala	Ala	Arg	Ala	Gly	Ser	Ala	Gly	Gln	Ser	Leu	Leu	Arg	Gln	
				20					25					30	
Leu	Pro	His	Arg	Cys	Pro	Pro	Leu	Arg	Pro	Arg	His	Arg	Arg	Pro	
				35					40					45	
Thr	Ala	Cys	Cys	Cys	Cys	Ser	Ala	Ala	Phe	Arg	Pro	Ala	Ala	Val	
				50					55					60	
Ala	Gln	Ala	Arg	Pro	Leu	Leu	Leu	Val	Leu	Gly	Ala	Val	Ala	Gln	
				65					70					75	
Gly	Arg	Ala	Val	Leu	His	Gln	Leu	Leu	Pro	Leu	Gln	Arg	Val	Gln	
				80					85					90	
Leu	Gln	Ile	His	Ser	Arg	Arg	Pro	Ser	Pro	Ile	Glu	Met	Pro	Gln	
				95					100					105	
Gly	Pro	His	Arg	Gly	Leu	Gly	Ala	Ala							
				110											

&lt;210&gt; 358

&lt;211&gt; 127

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1046117.1.orf1:2000SEP08

&lt;400&gt; 358

Leu	Cys	Leu	Thr	Cys	Leu	Leu	Glu	Gly	Asn	Thr	Gly	Lys	Pro	Gly	
1				5					10					15	
Val	Ala	Val	Thr	Leu	Val	Thr	Asn	Met	Ser	Gln	Asp	Ser	Val	Thr	
				20					25					30	
Phe	Ala	Asp	Val	Ala	Val	Asn	Phe	Thr	Lys	Glu	Glu	Trp	Thr	Leu	
				35					40					45	
Leu	Asp	Pro	Ala	Gln	Arg	Asn	Leu	Tyr	Arg	Asp	Val	Met	Leu	Glu	
				50					55					60	
Asn	Ser	Arg	Asn	Leu	Ala	Phe	Ile	Asp	Trp	Ala	Thr	Pro	Cys	Lys	

	65		70		75
Thr Lys Asp Ala	Thr Pro Gln Pro Asp	Ile Leu Pro Lys Arg	Thr		
	80		85		90
Phe Pro Glu Ala	Asn Arg Val Cys Leu Thr	Ser Ile Arg Phe	Pro		
	95		100		105
Ala Leu His Ile	Lys Arg Arg Leu Glu Met	Pro Gln Asn Arg	Gly		
	110		115		120
Thr Thr Gln Ala	Gly Gly Glu				
	125				

<210> 359  
 <211> 78  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:801015.1.orf1:2000SEP08

<400> 359	
Gly Ser Arg Lys Met Asp Ser Val Ala Phe Glu Asp Val Ala Val	
1 5 10 15	
Asn Phe Thr Gln Glu Glu Trp Ala Leu Leu Asp Pro Trp Gln Lys	
20 25 30	
Lys Leu Tyr Arg Asp Val Met Leu Glu Thr Tyr Arg Asn Leu Ala	
35 40 45	
Ser Val Gly Asp Asp Asp Asn Ile Pro Ser Leu Arg Glu Gln Val	
50 55 60	
Ala His Gln Arg Tyr Phe Lys Thr Trp His Val Glu Arg Glu Tyr	
65 70 75	
Phe Ser Lys	

<210> 360  
 <211> 158  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1175590.1.orf3:2000SEP08

<220>  
 <221> unsure  
 <222> 96  
 <223> unknown or other

<400> 360	
Ser Cys Pro Lys Ser Phe Arg Ala His Val Met Met His Ala Gly	
1 5 10 15	
Gly Arg Pro Tyr Glu Cys Lys His Cys Gly Lys Ala Phe Arg Cys	
20 25 30	
Gln Lys Ser Phe Arg Val His Met Ile Met His Ala Gly Gly Arg	
35 40 45	
Pro Tyr Glu Cys Lys Gln Cys Gly Lys Ala Tyr Cys Trp Ala Thr	
50 55 60	
Ser Phe Gln Arg His Val Arg Ile His Asn Gly Glu Lys Pro Tyr	
65 70 75	
Lys Cys Gly Lys Cys Gly Lys Ala Phe Gly Trp Pro Ser Ser Leu	
80 85 90	
His Lys His Ala Arg Xaa His Ala Arg Lys Lys Pro Val Ser Gly	
95 100 105	
Gly Ser Val Gly Lys Ser Ser Arg Glu Ala Leu Ala Pro Pro Gln	

	110		115		120
Met Ser Asn His Lys	Leu Glu Arg Lys	Ser Ile Asn Val Lys	Arg		
	125		130		135
Val Gly Lys Arg Met	Val Gly Pro His	Leu Tyr Thr Asn Met	Arg		
	140		145		150
Glu Ser Thr Leu Gly	Arg Asn Leu				
	155				

&lt;210&gt; 361

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1170585.2.orf2:2000SEP08

&lt;400&gt; 361

Asn Phe His Ile Gly	Ala Ser Arg Glu Asp	Leu Trp Ser Ile His	
1	5	10	15
Asp Leu Glu Ala Arg	Tyr Gln Glu Ser Gln	Ala Gly Asn Ser Arg	
	20	25	30
Asn Gly Glu Leu Thr	Lys His Gln Lys Thr	His Thr Thr Glu Lys	
	35	40	45
Ala Cys Glu Cys Lys	Glu Cys Gly Lys Phe	Phe Cys Gln Lys Ser	
	50	55	60
Ala Leu Ile Val His	Gln His Thr His Ser	Lys Gly Lys Ser Tyr	
	65	70	75
Asp Cys Asp Lys Cys	Gly Lys Ser Phe Ser	Lys Asn Glu Asp Leu	
	80	85	90
Ile Arg His Gln Lys	Ile His Thr Arg Asp	Lys Thr Tyr Glu Cys	
	95	100	105
Lys Glu Cys Lys Lys	Ile Phe Tyr His Gln		
	110	115	

&lt;210&gt; 362

&lt;211&gt; 59

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:719531.2.orf1:2000SEP08

&lt;400&gt; 362

Gln Gln Leu Phe Ile	Met Asn Thr Gly Cys	Arg Gln Thr His Thr	
1	5	10	15
Ala Pro Ala Thr Lys	Lys Phe Ala Gln Gly	His His Ser Leu Val	
	20	25	30
Leu Gly Ser Leu Lys	Leu Tyr Thr Gly Lys	Ser Ser Thr Tyr Cys	
	35	40	45
Ser Lys Asp Ala Val	Ser Gln Ala Cys Cys	Glu Pro Asn Gly	
	50	55	

&lt;210&gt; 363

&lt;211&gt; 64

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:794623.1.orf2:2000SEP08

&lt;400&gt; 363

Arg	Leu	Leu	Val	Arg	Glu	Lys	Glu	Thr	Gln	Lys	Arg	Lys	Arg	Lys	
1				5					10						15
Ala	Lys	Glu	Ser	Gly	Met	Ala	Leu	Pro	Gln	Gly	Leu	Leu	Thr	Phe	
				20					25						30
Arg	Asp	Val	Ala	Ile	Glu	Phe	Ser	Gln	Glu	Glu	Trp	Lys	Cys	Leu	
				35					40						45
Asp	Pro	Ala	Gln	Arg	Thr	Leu	Tyr	Arg	Asp	Val	Met	Leu	Glu	Asn	
				50					55						60
Tyr	Arg	Asn	Leu												

&lt;210&gt; 364

&lt;211&gt; 255

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1173119.1.orf3:2000SEP08

&lt;400&gt; 364

Glu	Gln	Gly	Leu	Tyr	Thr	Cys	Pro	Ala	His	Leu	His	Gln	His	Gln	
1				5					10						15
Lys	Glu	Gln	Ile	Arg	Glu	Lys	Leu	Ser	Arg	Gly	Asp	Gly	Gly	Arg	
				20					25						30
Pro	Thr	Phe	Val	Lys	Asn	His	Arg	Val	His	Met	Ala	Gly	Lys	Thr	
				35					40						45
Phe	Leu	Cys	Ser	Glu	Cys	Gly	Lys	Ala	Phe	Ser	His	Lys	His	Lys	
				50					55						60
Leu	Ser	Asp	His	Gln	Lys	Ile	His	Thr	Gly	Glu	Arg	Thr	Tyr	Lys	
				65					70						75
Cys	Ser	Lys	Cys	Gly	Ile	Leu	Phe	Met	Glu	Arg	Ser	Thr	Leu	Asn	
				80					85						90
Arg	His	Gln	Arg	Thr	His	Thr	Gly	Glu	Arg	Pro	Tyr	Glu	Cys	Asn	
				95					100						105
Glu	Cys	Gly	Lys	Ala	Phe	Leu	Cys	Lys	Ser	His	Leu	Val	Arg	His	
				110					115						120
Gln	Thr	Ile	His	Ser	Gly	Glu	Arg	Pro	Tyr	Glu	Cys	Ser	Glu	Cys	
				125					130						135
Gly	Lys	Leu	Phe	Met	Trp	Ser	Ser	Thr	Leu	Ile	Thr	His	Gln	Arg	
				140					145						150
Val	His	Thr	Gly	Lys	Arg	Pro	Tyr	Gly	Cys	Ser	Glu	Cys	Gly	Lys	
				155					160						165
Phe	Phe	Lys	Cys	Asn	Ser	Asn	Leu	Phe	Arg	His	Tyr	Arg	Ile	His	
				170					175						180
Thr	Gly	Lys	Arg	Ser	Tyr	Gly	Cys	Ser	Glu	Cys	Gly	Lys	Phe	Phe	
				185					190						195
Met	Glu	Arg	Ser	Thr	Leu	Ser	Arg	His	Gln	Arg	Val	His	Thr	Gly	
				200					205						210
Glu	Arg	Pro	Tyr	Glu	Cys	Asn	Glu	Cys	Gly	Lys	Phe	Phe	Ser	Leu	
				215					220						225
Lys	Ser	Val	Leu	Ile	Gln	His	Gln	Arg	Val	His	Thr	Gly	Glu	Arg	
				230					235						240
Pro	Tyr	Asp	Ala	Val	Ser	Val	Ala	Asn	Pro	Phe	Ala	Arg	Tyr	Tyr	
				245					250						255

&lt;210&gt; 365

&lt;211&gt; 97

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1093285.1.orf1:2000SEP08

&lt;400&gt; 365

Ala	Phe	Thr	Ser	Ala	His	Leu	Cys	Asn	Arg	Asp	Asn	Gly	Glu	Lys	
1				5					10					15	
Pro	Tyr	Lys	Cys	Ser	Glu	Cys	Gly	Lys	Ala	Phe	Arg	His	Lys	Leu	
				20					25					30	
Ser	Leu	Thr	Asn	His	Gln	Arg	Ile	His	Thr	Gly	Glu	Arg	Pro	Tyr	
				35					40					45	
Lys	Cys	Asn	Glu	Cys	Gly	Lys	Val	Phe	Asn	Arg	Ile	Ala	His	Leu	
				50					55					60	
Ala	Arg	His	Arg	Lys	Ile	His	Thr	Gly	Glu	Lys	Pro	Tyr	Lys	Cys	
				65					70					75	
Asn	Glu	Cys	Gly	Lys	Ala	Phe	Ser	Arg	Ile	Ser	Tyr	Leu	Ala	Gln	
				80					85					90	
His	Trp	Thr	Ile	His	Met	Gly									
				95											

&lt;210&gt; 366

&lt;211&gt; 158

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1091881.1.orf1:2000SEP08

&lt;400&gt; 366

Ala	Gly	Phe	Pro	Ser	Ala	Ala	Ser	Arg	Arg	Val	Pro	Gly	Asp	Cys	
1				5					10					15	
Arg	Arg	Leu	His	Arg	Pro	Val	Ala	Cys	Arg	Tyr	Cys	Glu	Ile	Tyr	
				20					25					30	
Arg	Glu	Asp	Ala	Gly	Thr	Pro	Lys	Ser	Trp	Glu	Met	Gly	Leu	Leu	
				35					40					45	
Ala	Phe	Arg	Asp	Val	Ala	Leu	Glu	Phe	Ser	Pro	Glu	Glu	Trp	Glu	
				50					55					60	
Cys	Leu	Asp	Pro	Ala	Gln	Arg	Ser	Leu	Tyr	Arg	Asp	Val	Met	Leu	
				65					70					75	
Glu	Asn	Tyr	Arg	Asn	Leu	Ile	Ser	Leu	Gly	Leu	Ala	Met	Ser	Lys	
				80					85					90	
Pro	Glu	Leu	Ile	Ile	Cys	Leu	Glu	Ala	Arg	Lys	Glu	Pro	Trp	Asn	
				95					100					105	
Val	Asn	Thr	Glu	Lys	Thr	Ala	Lys	His	Ser	Ala	Leu	Ser	Ser	Tyr	
				110					115					120	
Leu	Thr	Glu	Asp	Ile	Leu	Pro	Glu	Gln	Gly	Leu	Gln	Val	Ser	Phe	
				125					130					135	
Gln	Lys	Val	Met	Leu	Arg	Arg	Tyr	Glu	Arg	Cys	Cys	Leu	Glu	Lys	
				140					145					150	
Leu	Arg	Leu	Arg	Asn	Asp	Trp	Glu								
				155											

&lt;210&gt; 367

&lt;211&gt; 122

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1091617.1.orf2:2000SEP08

&lt;400&gt; 367

```

Asn Gln Thr Leu Leu Asn Ile Glu Gly Leu Thr His Gly Glu Lys
 1          5          10          15
Pro Tyr Lys Cys Asn Glu Cys Trp Arg Ser Phe Tyr Val Lys Ser
          20          25          30
Asn Leu Val Val His Gln Arg Asn Gln Gly Glu Lys Ser Tyr Arg
          35          40          45
Cys Pro Glu Cys Gly Lys Thr Phe Tyr Glu Lys Ser Ala Leu Thr
          50          55          60
Lys His Glu Arg Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Asn
          65          70          75
Glu Cys Arg Lys Thr Phe Ser Gln Arg Ser Ala Leu Thr Lys His
          80          85          90
Gln Arg Lys Thr His Lys Lys Lys Thr Ile Ile Asn Thr Leu His
          95          100          105
Val Gln Lys Pro Ala Ser Ser Arg Gln Ile Cys Gln Thr Ser Ala
          110          115          120
Lys Ser

```

&lt;210&gt; 368

&lt;211&gt; 242

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1082344.1.orf1:2000SEP08

&lt;400&gt; 368

```

Gly Glu Ala Ile His Gln Met Pro Asp Leu Thr Leu His Lys Lys
 1          5          10          15
Val Ser Ala Gly Glu Lys Pro Tyr Glu Cys Thr Lys Cys Arg Thr
          20          25          30
Val Phe Thr His Leu Ser Ser Leu Lys Arg His Val Lys Ser His
          35          40          45
Cys Gly Arg Lys Ala Pro Pro Gly Glu Glu Cys Lys Gln Ala Cys
          50          55          60
Ile Cys Pro Ser His Leu His Ser His Gly Arg Thr Asp Thr Glu
          65          70          75
Glu Lys Pro Tyr Lys Cys Gln Ala Cys Gly Gln Thr Phe Gln His
          80          85          90
Pro Arg Tyr Leu Ser His His Val Lys Thr His Thr Ala Glu Lys
          95          100          105
Thr Tyr Lys Cys Glu Gln Cys Arg Met Ala Phe Asn Gly Phe Ala
          110          115          120
Ser Phe Thr Arg His Val Arg Thr His Thr Lys Asp Arg Pro Tyr
          125          130          135
Lys Cys Gln Glu Cys Gly Arg Ala Phe Ile Tyr Pro Ser Thr Phe
          140          145          150
Gln Arg His Met Thr Thr His Thr Gly Glu Lys Pro Tyr Lys Cys
          155          160          165
Gln His Cys Gly Lys Ala Phe Thr Tyr Pro Gln Ala Phe Gln Arg
          170          175          180
His Glu Lys Thr His Thr Gly Glu Lys Pro His Lys Cys Lys Gln
          185          190          195
Cys Gly Met Ser Phe Lys Trp His Ser Ser Phe Arg Asn His Leu
          200          205          210
Arg Met His Thr Gly Gln Lys Ser His Glu Cys Gln Ser Tyr Ser
          215          220          225
Lys Ala Phe Ser Cys Gln Val Ile Leu Ser Lys Thr Ser Glu Ser
          230          235          240
Thr His

```

<210> 369  
 <211> 92  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1166249.1.orf3:2000SEP08

<400> 369  
 Gly Arg Thr Leu Trp Pro Cys Asp Leu Pro Pro Gly Ser Leu Glu  
 1 5 10 15  
 Met Gly Leu Leu Thr Phe Arg Asp Val Ala Ile Glu Phe Ser Leu  
 20 25 30  
 Glu Glu Trp Gln Cys Leu Asp Thr Ala Gln Lys Asn Leu Tyr Arg  
 35 40 45  
 Asn Val Met Leu Glu Asn Tyr Arg Asn Leu Ala Phe Leu Gly Ile  
 50 55 60  
 Ala Val Ser Lys Pro Asp Leu Ile Ile Cys Leu Glu Lys Glu Lys  
 65 70 75  
 Glu Pro Trp Asn Met Lys Arg Asp Glu Met Val Asp Glu Pro Pro  
 80 85 90  
 Gly Arg

<210> 370  
 <211> 85  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:799675.1.orf1:2000SEP08

<400> 370  
 Gly Arg His Gln Glu Gly Lys Glu Lys Glu Pro Gly Met Ala Leu  
 1 5 10 15  
 Pro Gln Gly Tyr Leu Thr Phe Arg Asp Val Ala Ile Glu Phe Ser  
 20 25 30  
 Leu Leu Glu Trp Lys Arg Leu Asp Pro Ala Gln Asn Ala Leu Tyr  
 35 40 45  
 Arg Ala Val Met Trp Glu Asn Tyr Arg Asn Leu Glu Ser Val Gly  
 50 55 60  
 Glu Glu Asn Val Pro Pro Asp Met Lys Asn Leu Ser Leu Cys Ile  
 65 70 75  
 Leu Ala Phe Pro Gly Phe Val Ser Leu Leu  
 80 85

<210> 371  
 <211> 118  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1178899.1.orf2:2000SEP08

<400> 371  
 Gly Cys Arg Gly Pro Pro Arg Arg Thr Ser Asp Leu Leu Gly Arg  
 1 5 10 15  
 Asp Leu Ile Arg Asp Leu Glu His Gly Trp Glu Glu Gln Arg Leu  
 20 25 30  
 Asp Leu Cys Pro Val Gln Thr Pro Thr Gly Trp Val Met Val Gly

	35		40		45									
Asp	Leu	Cys	Pro	Phe	Arg	Cys	Gly	Gly	Arg	Gly	Gly	Gln	Gly	Arg
	50								55					60
Arg	Ser	Arg	Gln	His	Arg	Phe	Leu	Gly	Ser	Cys	Glu	Gly	Ile	Met
	65								70					75
Arg	Arg	Ala	Glu	Leu	Ser	Ser	Gln	Val	Glu	Asp	Ser	Thr	Leu	His
	80								85					90
Ala	Trp	Ile	Arg	Tyr	Ser	Leu	Val	Leu	Asp	Val	Glu	Leu	Leu	Ala
	95								100					105
Tyr	Cys	Ser	Thr	Thr	Arg	Asp	Val	Arg	Met	Pro	Pro	Ser		
	110								115					

&lt;210&gt; 372

&lt;211&gt; 206

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1169241.1.orf3:2000SEP08

&lt;400&gt; 372

Glu	Thr	Arg	Pro	Glu	Met	Lys	Glu	Leu	Asp	Pro	Lys	Asn	Asp	Ile
1				5					10					15
Ser	Glu	Asp	Lys	Leu	Ser	Val	Val	Gly	Glu	Ala	Thr	Gly	Gly	Pro
				20					25					30
Thr	Arg	Asn	Gly	Ala	Arg	Gly	Pro	Gly	Ser	Glu	Gly	Val	Trp	Glu
				35					40					45
Pro	Gly	Ser	Trp	Pro	Glu	Arg	Pro	Arg	Gly	Asp	Ala	Gly	Ala	Glu
				50					55					60
Trp	Glu	Pro	Leu	Gly	Ile	Pro	Gln	Gly	Asn	Lys	Leu	Leu	Gly	Gly
				65					70					75
Ser	Val	Pro	Ala	Cys	His	Glu	Leu	Lys	Ala	Phe	Ala	Asn	Gln	Gly
				80					85					90
Cys	Val	Leu	Val	Pro	Pro	Arg	Leu	Asp	Asp	Pro	Thr	Glu	Lys	Gly
				95					100					105
Ala	Cys	Pro	Pro	Val	Arg	Arg	Gly	Lys	Asn	Phe	Ser	Ser	Thr	Ser
				110					115					120
Asp	Leu	Ser	Lys	Pro	Pro	Met	Pro	Cys	Glu	Glu	Lys	Lys	Thr	Tyr
				125					130					135
Asp	Cys	Ser	Glu	Cys	Gly	Lys	Ala	Phe	Ser	Arg	Ser	Ser	Ser	Leu
				140					145					150
Ile	Lys	His	Gln	Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Phe	Lys	Cys
				155					160					165
Asp	Thr	Cys	Gly	Lys	His	Ser	Ser	Ser	Ala	Arg	Pro	Ser	Pro	Ser
				170					175					180
Thr	Gly	Ala	Cys	Thr	Arg	Ala	Arg	Ser	Leu	Met	Arg	Ala	Pro	Ser
				185					190					195
Ala	Ala	Arg	Pro	Ser	Ala	Ser	Ala	Pro	Thr	Tyr				
				200					205					

&lt;210&gt; 373

&lt;211&gt; 206

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1180090.1.orf1:2000SEP08

&lt;400&gt; 373

Pro	His	Leu	Gly	Ser	Leu	Phe	Arg	Leu	Ser	Leu	Pro	Arg	Met	Ser
1				5					10					15



```

Gly Glu Ala Thr Val Leu Ala Tyr His Ala Pro Glu Glu Gln Glu
      20      25
Gly Leu Leu Val Val Lys Val Glu Glu Glu Asn Tyr Val Leu Asp
      35      40
Gln Asp Phe Gly Leu Gln Glu Asn Pro Trp Ser Gln Glu Val Phe
      50      55
Arg Gln Lys Phe Arg Gln Phe Ser Tyr Ser Asp Ser Thr Gly Pro
      65      70
Arg Glu Ala Leu Ser Arg Leu Arg Glu Leu Cys Cys Gln Trp Leu
      80      85
Arg Pro Glu Val His Ser Lys Glu Gln Ile Leu Glu Leu Leu Met
      95     100
Leu Glu Gln Phe Leu Ala Ile Leu Pro Glu Glu Leu Gln Ala Trp
     110     115
Leu Arg Glu His Arg Pro Glu Asn Gly Glu Glu Ala Val Thr Met
     125     130
Leu Glu Glu Leu Glu Lys Glu Leu Glu Glu Pro Arg Gln Gln Asp
     140     145
Thr Thr His Gly Gln Glu Met Phe Trp Gln Glu Met Thr Ser Thr
     155     160
Gly Ala Leu Lys Ser Leu Ser Leu Asn Ser Pro Val Gln Pro Leu
     170     175
Glu Asn Gln Cys Lys Thr Glu Thr Gln Glu Ser Gln Ala Phe Gln
     185     190
Glu Arg Gly Glu Lys Pro Gln Ser Met Trp Gly
     200     205

```

&lt;210&gt; 374

&lt;211&gt; 160

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2049322.1.orf1:2000SEP08

&lt;400&gt; 374

```

Arg Val Met Leu Cys Asp Glu Lys Ala Gln Lys Arg Arg Lys Arg
  1      5      10
Lys Ala Lys Glu Ser Gly Met Ala Leu Pro Gln Gly His Leu Thr
     20      25
Phe Arg Asp Val Ala Ile Glu Phe Ser Gln Ala Glu Trp Lys Cys
     35      40
Leu Asp Pro Ala Gln Arg Ala Leu Tyr Lys Asp Val Met Leu Glu
     50      55
Asn Tyr Arg Asn Leu Val Ser Leu Gly Ile Ser Leu Pro Asp Leu
     65      70
Asn Ile Asn Ser Met Leu Glu Gln Arg Arg Glu Pro Trp Ser Gly
     80      85
Glu Ser Glu Val Lys Ile Ala Lys Asn Ser Asp Gly Arg Glu Cys
     95     100
Ile Lys Gly Val Asn Thr Gly Ser Ser Tyr Ala Leu Gly Ser Asn
    110     115
Ala Glu Asp Lys Pro Ile Lys Lys Gln Leu Gly Val Ser Phe His
    125     130
Leu His Leu Ser Glu Leu Glu Leu Phe Pro Asp Glu Arg Val Ile
    140     145
Asn Gly Cys Asn Gln Val Glu Asn Phe Ile
    155     160

```

&lt;210&gt; 375

&lt;211&gt; 115

&lt;212&gt; PRT

Pro Ala Gly Ile Gly Arg Ser Thr Ala Lys Thr Pro Gly Thr Pro

1	5	10	15
Gly Ser Leu Glu Met Glu Asn Leu Lys Ser Gly Val Tyr Pro Leu			
	20	25	30
Lys Glu Ala Ser Gly Cys Pro Gly Ala Asp Arg Asn Leu Leu Val			
	35	40	45
Tyr Ser Phe Tyr Glu Lys Gly Pro Leu Thr Phe Arg Asp Val Ala			
	50	55	60
Ile Glu Phe Ser Leu Glu Glu Trp Gln Cys Leu Asp Thr Ala Gln			
	65	70	75
Gln Asp Leu Tyr Arg Lys Val Met Leu Glu Asn Tyr Arg Asn Leu			
	80	85	90
Val Phe Leu Ala Gly Ile Ala Val Ser Lys Pro Asp Leu Ile Thr			
	95	100	105
Cys Leu Glu Gln Gly Lys Glu Pro Trp Asn Met Lys Arg His Ala			
	110	115	120
Met Val Asp Gln Pro Pro Val Thr Tyr Ser His Phe Ala Gln Asp			
	125	130	135
Leu Trp Pro Glu Gln Gly Ile Lys Asp Ser Phe Gln Glu Val Ile			
	140	145	150
Leu Arg Arg Tyr Gly Lys Cys Gly His Glu Asp Leu Gln Leu Arg			
	155	160	165
Thr Gly Cys Lys Ser Val Asp Glu Cys Asn Leu His Lys Glu Cys			
	170	175	180
Tyr Asp Glu Leu Asn Gln Cys Leu Thr Thr Thr Gln Ser Glu Ile			
	185	190	195
Phe Gln Tyr Asp Lys Tyr Val Asn Val Phe Tyr Lys Phe Ser Asn			
	200	205	210
Pro Asn Ile Gln Lys Ile Arg His Thr Gly Lys Lys Pro Phe Lys			
	215	220	225
Cys Lys Lys Cys Asp Lys Ser Phe Cys Met Leu Leu His Leu Thr			
	230	235	240
Gln His Lys Arg Ile His Ile Arg Glu Asn Ser Tyr Gln Cys Glu			
	245	250	255
Glu Cys Gly Lys Val Phe Asn Trp Phe Ser Thr Leu Thr Arg His			
	260	265	270
Arg Ser Ala			

&lt;210&gt; 378

&lt;211&gt; 132

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1174107.2.orf3:2000SEP08

&lt;400&gt; 378

Pro Glu Asp Thr Gly Lys Ser Ile Ala Lys Met Pro Gly Pro Pro			
1	5	10	15
Glu Ser Leu Asp Met Gly Pro Leu Thr Phe Arg Asp Val Ala Ile			
	20	25	30
Glu Phe Ser Leu Glu Glu Trp Gln Cys Leu Asp Thr Ala Gln Gln			
	35	40	45
Asp Leu Tyr Arg Lys Val Met Leu Glu Asn Tyr Arg Asn Leu Val			
	50	55	60
Phe Leu Ala Gly Ile Ala Val Ser Lys Pro Asp Leu Val Thr Cys			
	65	70	75
Leu Glu Gln Gly Lys Asp Pro Trp Asn Met Lys Gly His Ser Thr			
	80	85	90
Val Val Lys Pro Pro Gly Phe Leu Thr Ala Ile Cys Asp Ser Phe			
	95	100	105
Leu Ile Cys Pro Lys Leu Tyr Val Leu Ile Leu Leu Lys Thr Phe			

	110		115		120
Ala	Gln	Gly	Gln	Ala	Leu
			Lys	Ile	Leu
				Phe	Lys
					Lys
	125		130		

<210> 379  
 <211> 233  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1177434.2.orf2:2000SEP08

<400> 379  
 Ala Pro Gly Asn Thr Pro Arg Gln Lys Pro Tyr Met Cys Val Leu  
   1                  5                  10                  15  
 Cys Gly Lys Gln Phe Trp Phe Ser Ala Asn Leu His Gln His Gln  
                   20                  25                  30  
 Lys Gln His Ser Gly Glu Lys Pro Phe Arg Ser Asp Lys Ser Arg  
                   35                  40                  45  
 Pro Phe Leu Leu Asn Asn Cys Ala Val Gln Ser Leu Glu Met Ser  
                   50                  55                  60  
 Phe Val Thr Gly Glu Ala Cys Lys Asp Phe Leu Ala Ser Ser Ser  
                   65                  70                  75  
 Ile Phe Glu His His Ala Pro His Asn Glu Trp Lys Pro His Ser  
                   80                  85                  90  
 Asn Thr Lys Cys Glu Glu Ala Ser His Cys Gly Lys Arg His Tyr  
                   95                  100                 105  
 Lys Cys Ser Glu Cys Gly Lys Thr Phe Ser Arg Lys Asp Ser Leu  
                  110                 115                 120  
 Val Gln His Gln Arg Val His Thr Gly Glu Arg Pro Tyr Glu Cys  
                  125                 130                 135  
 Gly Glu Cys Gly Lys Thr Phe Ser Arg Lys Pro Ile Leu Ala Gln  
                  140                 145                 150  
 His Gln Arg Ile His Thr Gly Glu Met Pro Tyr Glu Cys Gly Ile  
                  155                 160                 165  
 Cys Gly Lys Val Phe Asn His Ser Ser Asn Leu Ile Val His Gln  
                  170                 175                 180  
 Arg Val His Thr Gly Ala Arg Pro Tyr Lys Cys Ser Glu Cys Gly  
                  185                 190                 195  
 Lys Ala Tyr Ser His Lys Ser Thr Leu Val Gln His Glu Ser Ile  
                  200                 205                 210  
 His Thr Gly Glu Arg Pro Tyr Glu Cys Ser Glu Cys Gly Lys Tyr  
                  215                 220                 225  
 Ser Trp Ser Gln Ile Gln Thr His  
                  230

<210> 380  
 <211> 140  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1184255.1.orf2:2000SEP08

<400> 380  
 Lys Asn Phe Arg His Lys Phe Ser Leu Thr Asn His Gln Arg Ser  
   1                  5                  10                  15  
 His Thr Ala Glu Lys Pro Tyr Lys Cys Asn Glu Cys Gly Lys Val  
                   20                  25                  30  
 Phe Ser Leu Leu Ser Tyr Leu Ala Arg His Gln Ile Ile His Ser  
                   35                  40                  45

Thr	Glu	Lys	Pro	Tyr	Lys	Cys	Asn	Glu	Cys	Gly	Arg	Ala	Phe	His	
				50					55					60	
Lys	Arg	Pro	Gly	Leu	Met	Ala	His	Leu	Leu	Ile	His	Thr	Gly	Glu	
				65					70					75	
Lys	Pro	Tyr	Lys	Cys	Asn	Glu	Cys	Asp	Lys	Val	Phe	Gly	Arg	Lys	
				80					85					90	
Leu	Tyr	Leu	Thr	Asn	His	Gln	Arg	Ile	His	Thr	Gly	Glu	Arg	Pro	
				95					100					105	
Tyr	Lys	Cys	Asn	Ala	Cys	Gly	Lys	Val	Phe	Asn	Gln	Asn	Pro	His	
				110					115					120	
Leu	Ser	Arg	His	Arg	Lys	Ile	His	Ala	Gly	Glu	Asn	Ser	Leu	Arg	
				125					130					135	
Thr	Leu	Gln	Met	Glu											
				140											

&lt;210&gt; 381

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1164555.1.orf3:2000SEP08

&lt;400&gt; 381

Glu	Lys	Thr	Leu	Ala	Lys	Val	Met	Ser	Val	Ala	Lys	Pro	Phe	Asn	
1				5					10					15	
Tyr	Ser	Ser	Leu	Leu	Arg	Arg	His	His	Ile	Thr	His	Ser	Arg	Glu	
				20					25					30	
Arg	Glu	Tyr	Lys	Cys	Asp	Val	Cys	Gly	Lys	Ile	Phe	Asn	Gln	Lys	
				35					40					45	
Gln	Tyr	Ile	Val	Tyr	His	His	Arg	Cys	His	Thr	Gly	Glu	Lys	Thr	
				50					55					60	
Tyr	Lys	Cys	Asn	Glu	Cys	Gly	Lys	Thr	Phe	Thr	Gln	Met	Ser	Ser	
				65					70					75	
Leu	Val	Cys	His	Arg	Arg	Leu	His	Thr	Gly	Glu	Lys	Pro	Tyr	Lys	
				80					85					90	
Cys	Asn	Glu	Cys	Gly	Lys	Thr	Phe	Ser	Glu	Lys	Ser	Ser	Leu	Arg	
				95					100					105	
Cys	His	Arg	Arg	Leu	His	Thr	Gly	Glu	Lys	Pro	Tyr	Lys	Cys	Asn	
				110					115					120	
Glu	Cys	Gly	Lys	Thr	Phe	Gly	Arg	Asn	Ser	Ala	Leu	Val	Ile	His	
				125					130					135	
Lys	Ala	Arg	Ser	Ser	Leu	Cys	Gly	His	Ser	Met	Gln	Asn	Ile	Arg	
				140					145					150	
Lys	Phe	Ile													

&lt;210&gt; 382

&lt;211&gt; 420

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:238666.4.orf1:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 379, 406

&lt;223&gt; unknown or other

&lt;400&gt; 382

Gly	Pro	Phe	Gly	Gln	Pro	Leu	Arg	Arg	Ser	Glu	Leu	Glu	Thr	Thr	
1				5					10					15	
Glu	Pro	Ser	Ala	Gly	Ala	Gly	Gly	Ala	Ala	Arg	Gly	Ser	Gly	Trp	
				20					25					30	
Ala	Gly	Gly	Arg	Glu	Ile	Arg	Ser	Pro	Pro	Glu	Ala	Val	Ile	Leu	
				35					40					45	
Thr	Pro	Ala	Pro	Glu	Pro	Gly	Pro	Gly	Glu	Ala	Gly	Met	Ala	Pro	
				50					55					60	
Arg	Pro	Pro	Thr	Ala	Ala	Pro	Gln	Glu	Ser	Val	Thr	Phe	Lys	Asp	
				65					70					75	
Val	Ser	Val	Asp	Phe	Thr	Gln	Glu	Glu	Trp	Tyr	His	Val	Asp	Pro	
				80					85					90	
Ala	Gln	Arg	Ser	Leu	Tyr	Arg	Asp	Val	Met	Leu	Glu	Asn	Tyr	Ser	
				95					100					105	
His	Leu	Val	Ser	Leu	Gly	Tyr	Gln	Val	Ser	Lys	Pro	Glu	Val	Ile	
				110					115					120	
Phe	Lys	Leu	Glu	Gln	Gly	Glu	Glu	Pro	Trp	Ile	Ser	Glu	Gly	Glu	
				125					130					135	
Ile	Gln	Arg	Pro	Phe	Tyr	Pro	Asp	Trp	Lys	Thr	Arg	Pro	Glu	Val	
				140					145					150	
Lys	Ser	Ser	His	Leu	Gln	Gln	Asp	Val	Ser	Glu	Val	Ser	His	Cys	
				155					160					165	
Thr	His	Asp	Leu	Leu	His	Ala	Thr	Leu	Glu	Asp	Ser	Trp	Asp	Val	
				170					175					180	
Ser	Ser	Gln	Leu	Asp	Arg	Gln	Gln	Glu	Asn	Trp	Lys	Arg	His	Leu	
				185					190					195	
Gly	Ser	Glu	Ala	Ser	Thr	Gln	Lys	Lys	Ile	Ile	Thr	Pro	Gln	Glu	
				200					205					210	
Asn	Phe	Glu	Gln	Asn	Lys	Phe	Gly	Glu	Asn	Ser	Arg	Leu	Asn	Thr	
				215					220					225	
Asn	Leu	Val	Thr	Gln	Leu	Asn	Ile	Pro	Ala	Arg	Ile	Arg	Pro	Ser	
				230					235					240	
Glu	Cys	Glu	Thr	Leu	Gly	Ser	Asn	Leu	Gly	His	Asn	Ala	Asp	Leu	
				245					250					255	
Leu	Asn	Glu	Asn	Asn	Ile	Leu	Ala	Lys	Lys	Lys	Pro	Tyr	Lys	Cys	
				260					265					270	
Asp	Lys	Cys	Arg	Lys	Ala	Phe	Ile	His	Arg	Ser	Ser	Leu	Thr	Lys	
				275					280					285	
His	Glu	Lys	Thr	His	Lys	Gly	Glu	Gly	Ala	Phe	Pro	Asn	Gly	Thr	
				290					295					300	
Asp	Gln	Gly	Ile	Tyr	Pro	Gly	Lys	Lys	His	His	Glu	Cys	Thr	Asp	
				305					310					315	
Cys	Gly	Lys	Thr	Phe	Leu	Trp	Lys	Thr	Gln	Leu	Thr	Glu	His	Gln	
				320					325					330	
Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Phe	Glu	Cys	Asn	Val	Cys	Gly	
				335					340					345	
Lys	Ala	Phe	Arg	His	Ser	Ser	Ser	Leu	Gly	Gln	His	Glu	Asn	Ala	
				350					355					360	
His	Thr	Gly	Glu	Lys	Pro	Tyr	Gln	Cys	Ser	Leu	Cys	Gly	Lys	Ala	
				365					370					375	
Phe	Gln	Arg	Xaa	Pro	Pro	Leu	Phe	Asn	Thr	Ser	Glu	Phe	Thr	Leu	
				380					385					390	
Glu	Arg	Asn	Pro	Ile	Asp	Val	Asn	Leu	Cys	Gly	Arg	Ser	Phe	Arg	
				395					400					405	
Xaa	Gly	Thr	Ser	Leu	Thr	Gln	His	Glu	Gly	His	Thr	Gln	Trp	Arg	
				410					415					420	

&lt;210&gt; 383

&lt;211&gt; 68

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1166752.1.orf1:2000SEP08

&lt;400&gt; 383

```

Phe Leu Leu Trp Arg Val Ser Cys Ile Leu Trp Gln Tyr Arg Phe
 1             5             10             15
Pro His Ser Leu His Thr Phe Leu Val His Phe Ile Thr Val Ser
             20             25             30
Asn Cys Asp Asn Thr Ser Leu Ile Ala Leu Phe Ala Asp Asn Ser
             35             40             45
Leu Ile Ser Ala Pro Gly Thr Ser Glu Arg Gly Lys Lys Ile Asn
             50             55             60
Glu Gln Lys Gln Ile Lys Leu Met
             65

```

&lt;210&gt; 384

&lt;211&gt; 186

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2049654.1.orf3:2000SEP08

&lt;400&gt; 384

```

Asp Val Ala Leu Ala Ile Ser Arg Arg Arg Gln Arg Leu Leu Glu
 1             5             10             15
Glu Arg His Gly Ala Asp Gln Ala Arg Trp Ser Lys Pro Gln Arg
             20             25             30
Glu Ser Gly Ser Gly Arg Glu Ile Val Trp Gly Glu Gly Arg Arg
             35             40             45
Gln Ser Arg Gly Gly Ala Gly Cys Gly Pro Phe Trp Arg Thr Leu
             50             55             60
Ile Glu Cys Pro Lys Leu Pro Gly Met Cys Val Asp Pro Ser Ser
             65             70             75
Leu Arg Ser Pro Ala Gly Val Lys Leu His Thr Cys Leu His Ser
             80             85             90
Thr Leu Pro Val Pro Lys Arg Ala His Arg Arg Arg Lys Asn Gly
             95             100            105
Cys Gln Pro Pro Ala Asn His Gly Pro Gly Ile Gly Asp Leu Gln
             110            115            120
Gly Cys Gly Tyr Thr Val His Pro Gly Arg Val Gly Ala Ala Glu
             125            130            135
Pro Arg Pro Glu Gly Pro Val Gln Gly Arg Asp Ala Gly Glu Leu
             140            145            150
Gln Gln Pro Gly Leu Thr Gly Thr Leu Arg Thr Gln Thr Arg Tyr
             155            160            165
Val Phe Pro Ala Arg Lys Lys Gly Ser Val Asp Ala Arg Gly His
             170            175            180
Pro Trp Arg Leu Leu Ser
             185

```

&lt;210&gt; 385

&lt;211&gt; 81

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:242665.2.orf2:2000SEP08

&lt;400&gt; 385

```

Leu Leu Asp Pro Ser Thr Asp Cys Gly Phe Trp Val Cys Leu Leu
 1          5          10
Leu Val Ile Gly Pro Arg Glu Asn Ala Ala Arg Pro Gly Leu Val
          20          25          30
Pro Gly Thr Ile Arg Ile Lys Thr Cys Trp Val Pro Thr Thr His
          35          40          45
Val Val Val Ser Ser Leu His Thr Leu Ser Leu Glu His Trp Asp
          50          55          60
Ala Ala Trp Ser Gln Glu Asp Asp Ser Val Val His Ser Ser Ala
          65          70          75
Leu Val Met His Tyr Ala
          80

```

&lt;210&gt; 386

&lt;211&gt; 476

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:208637.1.orf1:2000SEP08

&lt;400&gt; 386

```

Arg Arg Gly Arg Ile Ala Gly Glu Cys Gly Phe Phe Leu Leu Gly
 1          5          10          15
Asp Ser Gln Pro Ser Ala Arg Pro Glu Glu Glu Lys Arg Arg Arg
          20          25          30
Arg Glu Ala Gly Met Glu Ser Thr Lys Leu Ile Ser Ala Thr Asp
          35          40          45
Ile Gln Tyr Ser Gly Ser Leu Leu Asn Ser Leu Asn Glu Gln Arg
          50          55          60
Gly His Gly Leu Phe Cys Asp Val Thr Val Ile Val Glu Asp Arg
          65          70          75
Lys Phe Arg Ala His Lys Asn Ile Leu Ser Ala Ser Ser Thr Tyr
          80          85          90
Phe His Gln Leu Phe Ser Val Ala Gly Gln Val Val Glu Leu Ser
          95          100          105
Phe Ile Arg Ala Glu Ile Phe Ala Glu Ile Leu Asn Tyr Ile Tyr
          110          115          120
Ser Ser Lys Ile Val Arg Val Arg Ser Asp Leu Leu Asp Glu Leu
          125          130          135
Ile Lys Ser Gly Gln Leu Leu Gly Val Lys Phe Ile Ala Glu Leu
          140          145          150
Gly Val Pro Leu Ser Gln Val Lys Ser Ile Ser Gly Thr Ala Gln
          155          160          165
Asp Gly Asn Thr Glu Pro Leu Pro Pro Asp Ser Gly Asp Lys Asn
          170          175          180
Leu Val Ile Gln Lys Ser Lys Gly Glu Ala Gln Asp Asn Gly Ala
          185          190          195
Thr Ile Met Pro Ile Ile Thr Glu Ser Phe Ser Leu Ser Ala Glu
          200          205          210
Asp Tyr Glu Met Lys Lys Ile Ile Val Thr Asp Ser Asp Asp Asp
          215          220          225
Asp Asp Asp Val Ile Phe Cys Ser Glu Ile Leu Pro Thr Lys Glu
          230          235          240
Thr Leu Pro Ser Asn Asn Thr Val Ala Gln Val Gln Ser Asn Pro
          245          250          255
Gly Pro Val Ala Ile Ser Asp Val Ala Pro Ser Ala Ser Asn Asn
          260          265          270
Ser Pro Pro Leu Thr Asn Ile Thr Pro Thr Gln Lys Leu Pro Thr
          275          280          285
Pro Val Asn Gln Ala Thr Leu Ser Gln Thr Gln Gly Ser Glu Lys
          290          295          300

```



Leu	Leu	Val	Ser	Ser	Ala	Pro	Thr	His	Leu	Thr	Pro	Asn	Ile	Ile	
				305					310					315	
Leu	Leu	Asn	Gln	Thr	Pro	Leu	Ser	Thr	Pro	Pro	Asn	Val	Ser	Ser	
				320					325					330	
Ser	Leu	Pro	Asn	His	Met	Pro	Ser	Ser	Ile	Asn	Leu	Leu	Val	Gln	
				335					340					345	
Asn	Gln	Gln	Thr	Pro	Asn	Ser	Ala	Ile	Leu	Thr	Gly	Asn	Lys	Ala	
				350					355					360	
Asn	Glu	Glu	Glu	Glu	Glu	Glu	Ile	Ile	Asp	Asp	Asp	Asp	Asp	Thr	
				365					370					375	
Ile	Ser	Ser	Ser	Pro	Asp	Ser	Ala	Val	Ser	Asn	Thr	Ser	Leu	Val	
				380					385					390	
Pro	Gln	Ala	Asp	Thr	Ser	Gln	Asn	Thr	Ser	Phe	Asp	Gly	Ser	Leu	
				395					400					405	
Ile	Gln	Lys	Met	Gln	Ile	Pro	Thr	Leu	Leu	Gln	Glu	Pro	Leu	Ser	
				410					415					420	
Asn	Ser	Leu	Lys	Ile	Ser	Asp	Ile	Ile	Thr	Arg	Asn	Thr	Asn	Asp	
				425					430					435	
Pro	Gly	Val	Gly	Ser	Lys	His	Leu	Met	Glu	Gly	Gln	Lys	Ile	Ile	
				440					445					450	
Thr	Leu	Asp	Thr	Ala	Thr	Glu	Ile	Glu	Gly	Leu	Ser	Thr	Gly	Cys	
				455					460					465	
Lys	Val	Tyr	Ala	Asn	Ile	Gly	Glu	Glu	Tyr	Leu					
				470					475						

&lt;210&gt; 387

&lt;211&gt; 292

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2051808.1.orf1:2000SEP08

&lt;400&gt; 387

Ala	Pro	Thr	Met	Ala	Ser	Lys	Cys	Pro	Lys	Cys	Asp	Lys	Thr	Val	
1				5					10					15	
Tyr	Phe	Ala	Glu	Lys	Val	Ser	Ser	Leu	Gly	Lys	Asp	Trp	His	Lys	
				20					25					30	
Phe	Cys	Leu	Lys	Cys	Glu	Arg	Cys	Asn	Lys	Thr	Leu	Thr	Pro	Gly	
				35					40					45	
Gly	His	Ala	Glu	His	Asp	Gly	Lys	Pro	Phe	Cys	His	Lys	Pro	Cys	
				50					55					60	
Tyr	Ala	Thr	Leu	Phe	Gly	Pro	Lys	Gly	Val	Asn	Ile	Gly	Gly	Ala	
				65					70					75	
Gly	Ser	Tyr	Ile	Tyr	Glu	Lys	Pro	Pro	Thr	Glu	Ala	Pro	Gln	Val	
				80					85					90	
Thr	Gly	Pro	Ile	Glu	Val	Pro	Val	Val	Arg	Thr	Glu	Glu	Arg	Lys	
				95					100					105	
Thr	Ser	Gly	Pro	Pro	Lys	Gly	Pro	Ser	Lys	Ala	Ser	Ser	Val	Thr	
				110					115					120	
Thr	Phe	Thr	Gly	Glu	Pro	Asn	Met	Cys	Pro	Arg	Cys	Asn	Lys	Arg	
				125					130					135	
Val	Tyr	Phe	Ala	Glu	Lys	Val	Thr	Ser	Leu	Gly	Lys	Asp	Trp	His	
				140					145					150	
Arg	Pro	Cys	Leu	Arg	Cys	Glu	Arg	Cys	Ser	Lys	Thr	Leu	Thr	Pro	
				155					160					165	
Gly	Gly	His	Ala	Glu	His	Asp	Gly	Gln	Pro	Tyr	Cys	His	Lys	Pro	
				170					175					180	
Cys	Tyr	Gly	Ile	Leu	Phe	Gly	Pro	Lys	Gly	Arg	Glu	Tyr	Trp	Cys	
				185					190					195	
Cys	Gly	Gln	Leu	His	Leu	Arg	Gln	Gly	Pro	Gly	Arg	His	Ser	Ser	
				200					205					210	

```
<210> 388
<211> 98
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: LI:1175136.1.orf2:2000SEP08
```

```
<210> 389
<211> 171
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: LI:1177337.1.orf1:2000SEP08
```

193/283

	110		115		120
Thr Leu Thr Lys	His Arg Lys Ile His	Thr Arg Gln Lys Pro Tyr			
	125		130		135
Asn Cys Glu Glu Cys Asp Asn Thr Phe	Asn Gln Ser Ser Asn Leu				
	140		145		150
Ile Lys Gln Asn Asn Ser Tyr Trp Arg	Glu Thr Leu Gln Met Ser				
	155		160		165
Arg Met Trp Glu Ser Leu					
	170				

&lt;210&gt; 390

&lt;211&gt; 336

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1165056.1.orf2:2000SEP08

&lt;400&gt; 390

Thr Val Met Leu Cys Asp Glu Glu Ala Gln Lys Arg Lys Ala Lys		
1	5	10
Glu Ser Gly Met Ala Leu Pro Gln Gly Arg Leu Thr Phe Met Asp		
	20	25
Val Ala Ile Glu Phe Ser Gln Glu Glu Trp Lys Ser Leu Asp Pro		
	35	40
Gly Gln Arg Ala Leu Tyr Arg Asp Val Met Leu Glu Asn Tyr Arg		
	50	55
Asn Leu Val Phe Leu Gly Ile Cys Leu Pro Asp Leu Ser Ile Ile		
	65	70
Ser Met Leu Lys Gln Arg Arg Glu Pro Leu Ile Leu Gln Ser Gln		
	80	85
Val Lys Ile Val Lys Asn Thr Asp Gly Arg Glu Cys Val Arg Ser		
	95	100
Val Asn Thr Gly Arg Ser Cys Val Leu Gly Ser Asn Ala Glu Asn		
	110	115
Lys Pro Ile Lys Asn Gln Leu Gly Leu Thr Leu Glu Ser His Leu		
	125	130
Ser Glu Leu Gln Leu Phe Gln Ala Gly Arg Lys Ile Tyr Arg Ser		
	140	145
Asn Gln Val Glu Lys Phe Thr Asn His Arg Ser Ser Val Ser Pro		
	155	160
Leu Gln Lys Ile Ser Ser Ser Phe Thr Thr His Ile Phe Asn Lys		
	170	175
Tyr Arg Asn Asp Leu Ile Asp Phe Pro Leu Leu Pro Gln Glu Glu		
	185	190
Lys Ala Tyr Ile Arg Gly Lys Ser Tyr Glu Tyr Glu Cys Ser Glu		
	200	205
Asp Gly Glu Val Phe Arg Val Arg Ala Ser Leu Thr Asn His Gln		
	215	220
Val Ile His Thr Ala Glu Lys Pro Tyr Lys Cys Thr Glu Cys Gly		
	230	235
Lys Val Phe Ser Arg Asn Ser His Leu Val Glu His Trp Arg Ile		
	245	250
His Thr Gly Gln Lys Pro Tyr Lys Cys Ser Glu Cys Asp Lys Val		
	260	265
Phe Asn Arg Asn Ser Asn Leu Ala Arg His Gln Arg Ile His Thr		
	275	280
Gly Glu Lys Pro His Lys Cys Asn Glu Cys Gly Lys Ala Phe Arg		
	290	295
Glu Cys Ser Gly Leu Thr Thr His Leu Val Ile His Thr Gly Glu		
	305	310
Lys Pro Tyr Lys Cys Asn Glu Cys Gly Lys Asn Phe Arg His Lys		

320  
Phe Ser Leu Thr Asn His  
335

325

330

<210> 391  
<211> 375  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: LI:1175250.1.orf1:2000SEP08

<400> 391  
Leu Thr Pro Gly His Pro Gly Ser Arg Gly Met Asp Ser Val Ala  
1 5 10 15  
Phe Glu Asp Val Ala Val Asn Phe Thr Gln Glu Glu Trp Ala Leu  
20 25 30  
Leu Asp Ser Ser Gln Lys Asn Leu Tyr Arg Glu Val Met Gln Glu  
35 40 45  
Thr Cys Arg Asn Leu Ala Ser Val Gly Ser Gln Trp Lys Asp Gln  
50 55 60  
Asn Ile Glu Asp His Phe Glu Lys Pro Gly Lys Asp Ile Arg Asn  
65 70 75  
His Ile Val Gln Arg Leu Cys Glu Ser Lys Glu Asp Gly Gln Tyr  
80 85 90  
Gly Glu Val Val Ser Gln Ile Pro Asn Leu Asp Leu Asn Glu Asn  
95 100 105  
Ile Ser Thr Gly Leu Lys Pro Cys Glu Cys Ser Ile Cys Gly Lys  
110 115 120  
Val Phe Val Arg His Ser Leu Leu Asn Arg His Ile Leu Ala His  
125 130 135  
Ser Gly Tyr Lys Pro Tyr Gly Glu Lys Gln Tyr Lys Cys Glu Gln  
140 145 150  
Cys Gly Lys Phe Phe Val Ser Val Pro Gly Val Arg Arg His Met  
155 160 165  
Ile Met His Ser Gly Asn Pro Ala Tyr Lys Cys Thr Ile Cys Gly  
170 175 180  
Lys Ala Phe Tyr Phe Leu Asn Ser Val Glu Arg His Gln Arg Thr  
185 190 195  
His Thr Gly Glu Lys Pro Tyr Lys Cys Lys Gln Cys Gly Lys Ala  
200 205 210  
Phe Thr Val Ser Gly Ser Cys Leu Ile His Glu Arg Thr His Thr  
215 220 225  
Gly Glu Lys Pro Tyr Glu Cys Lys Glu Cys Gly Lys Thr Phe Arg  
230 235 240  
Phe Ser Cys Ser Phe Lys Thr His Glu Arg Thr His Thr Gly Glu  
245 250 255  
Arg Pro Tyr Lys Cys Thr Lys Cys Asp Lys Ala Phe Ser Cys Ser  
260 265 270  
Thr Ser Leu Arg Tyr His Gly Ser Ile His Thr Gly Glu Arg Pro  
275 280 285  
Tyr Glu Cys Lys Gln Cys Gly Lys Ala Phe Ser Arg Leu Ser Ser  
290 295 300  
Leu Cys Asn His Arg Ser Thr His Thr Gly Glu Lys Pro Tyr Glu  
305 310 315  
Cys Lys Gln Cys Asp Gln Ala Phe Ser Arg Leu Ser Ser Leu His  
320 325 330  
Leu His Glu Arg Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Lys  
335 340 345  
Lys Cys Gly Lys Ala Tyr Thr Arg Ser Ser His Leu Thr Arg His  
350 355 360  
Glu Arg Ser His Asp Ile Glu Ala Gly Cys Ser Asp Ser Ala Tyr

365

370

375

&lt;210&gt; 392

&lt;211&gt; 444

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1183192.1.orf3:2000SEP08

&lt;400&gt; 392

Arg Gly Ser Arg Ser Gly Phe Thr Ser Pro Asn Pro Glu Ala Val	1	5	10	15
Cys Glu Ala Gly Thr Pro Ala Met Phe Gln Thr Ala Trp Arg Gln	20	25	30	35
Glu Pro Val Thr Phe Glu Asp Val Ala Val Tyr Phe Thr Gln Asn	40	45	50	55
Glu Trp Ala Ser Leu Asp Ser Val Gln Arg Ala Leu Tyr Arg Glu	60	65	70	75
Val Met Leu Glu Asn Tyr Ala Asn Val Ala Ser Leu Ala Phe Pro	80	85	90	95
Phe Thr Thr Pro Val Leu Val Ser Gln Leu Glu Gln Gly Glu Leu	100	105	110	115
Pro Trp Gly Leu Asp Pro Trp Glu Pro Met Gly Arg Glu Ala Leu	120	125	130	135
Arg Gly Ile Cys Pro Gly Asp Glu Ala Arg Thr Glu Lys Glu Gly	140	145	150	155
Leu Thr Pro Lys Asp His Val Ser Lys Glu Thr Glu Ser Phe Arg	160	165	170	175
Leu Met Val Gly Gly Leu Pro Gly Asn Val Ser Gln His Leu Asp	180	185	190	195
Phe Gly Ser Ser Leu Glu Gln Pro Gln Gly His Trp Ile Ile Lys	200	205	210	215
Thr Lys Ser Lys Arg Arg His Phe Thr Asp Thr Ser Ala Arg His	220	225	230	235
His Glu Ala Tyr Glu Val Lys Asn Gly Glu Lys Phe Glu Lys Leu	240	245	250	255
Gly Lys Asn Ile Ser Val Ser Thr Gln Leu Thr Thr Asn Gln Thr	260	265	270	275
Asn Pro Ser Gly Gln Ile Ser Tyr Glu Cys Gly Gln Cys Gly Arg	280	285	290	295
Tyr Phe Ile Gln Met Ala Asp Phe His Arg His Glu Lys Cys His	300	305	310	315
Thr Gly Glu Lys Ser Phe Glu Cys Lys Glu Cys Gly Lys Tyr Phe	320	325	330	335
Arg Tyr Asn Ser Leu Leu Ile Arg His Gln Ile Ile His Thr Gly	340	345	350	355
Lys Lys Pro Phe Lys Cys Lys Glu Cys Gly Lys Gly Leu Ser Ser	360	365	370	375
Asp Thr Ala Leu Ile Gln His Gln Arg Ile His Thr Gly Glu Lys				
Pro Tyr Glu Cys Lys Glu Cys Gly Lys Ala Phe Ser Ser Ser Ser				
Val Phe Leu Gln His Gln Arg Phe His Thr Gly Glu Lys Leu Tyr				
Glu Cys Asn Glu Cys Trp Lys Thr Phe Ser Cys Ser Ser Ser Phe				
Thr Val His Gln Arg Met His Thr Gly Glu Lys Pro Tyr Glu Cys				
Lys Glu Cys Gly Lys Arg Leu Ser Ser Asn Thr Ala Leu Thr Gln				

His	Gln	Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Phe	Glu	Cys	Lys	Glu
				380					385					390
Cys	Gly	Lys	Ala	Phe	Asn	Gln	Lys	Ile	Thr	Leu	Ile	Gln	His	Gln
				395					400					405
Arg	Val	His	Thr	Gly	Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Val	Cys	Gly
				410					415					420
Lys	Thr	Phe	Ser	Trp	Cys	Gly	Arg	Phe	Ile	Leu	His	Gln	Lys	Leu
				425					430					435
His	Thr	Gln	Lys	Thr	Pro	Val	Gln	Ala						
				440										

&lt;210&gt; 393

&lt;211&gt; 620

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1183325.1.orf1:2000SEP08

&lt;400&gt; 393

Gln	Lys	Lys	Met	Ile	Lys	Ser	Gln	Glu	Ser	Leu	Thr	Leu	Glu	Asp
1				5					10					15
Val	Ala	Val	Glu	Phe	Thr	Trp	Glu	Glu	Trp	Gln	Leu	Leu	Gly	Pro
				20					25					30
Ala	Gln	Lys	Asp	Leu	Tyr	Arg	Asp	Val	Met	Leu	Glu	Asn	Tyr	Ser
				35					40					45
Asn	Leu	Val	Ser	Val	Gly	Tyr	Gln	Ala	Ser	Lys	Pro	Asp	Ala	Leu
				50					55					60
Phe	Lys	Leu	Glu	Gln	Gly	Glu	Pro	Trp	Thr	Val	Glu	Asn	Glu	Ile
				65					70					75
His	Ser	Gln	Ile	Cys	Pro	Glu	Ile	Lys	Lys	Val	Asp	Asn	His	Leu
				80					85					90
Gln	Met	His	Ser	Gln	Lys	Gln	Arg	Cys	Leu	Lys	Arg	Val	Glu	Gln
				95					100					105
Cys	His	Lys	His	Asn	Ala	Phe	Gly	Asn	Ile	Ile	His	Gln	Arg	Lys
				110					115					120
Ser	Asp	Phe	Pro	Leu	Arg	Gln	Asn	His	Asp	Thr	Phe	Asp	Leu	His
				125					130					135
Gly	Lys	Ile	Leu	Lys	Ser	Asn	Leu	Ser	Leu	Val	Asn	Gln	Asn	Lys
				140					145					150
Arg	Tyr	Glu	Ile	Lys	Asn	Ser	Val	Gly	Val	Asn	Gly	Asp	Gly	Lys
				155					160					165
Ser	Phe	Leu	His	Ala	Lys	His	Glu	Gln	Phe	His	Asn	Glu	Met	Asn
				170					175					180
Phe	Pro	Glu	Gly	Gly	Asn	Ser	Val	Asn	Thr	Asn	Ser	Gln	Phe	Ile
				185					190					195
Lys	His	Gln	Arg	Thr	Gln	Asn	Ile	Asp	Lys	Pro	His	Val	Cys	Thr
				200					205					210
Glu	Cys	Gly	Lys	Ala	Phe	Leu	Lys	Lys	Ser	Arg	Leu	Ile	Tyr	His
				215					220					225
Gln	Arg	Val	His	Thr	Gly	Glu	Lys	Pro	His	Gly	Cys	Ser	Ile	Cys
				230					235					240
Gly	Lys	Ala	Phe	Ser	Arg	Lys	Ser	Gly	Leu	Thr	Glu	His	Gln	Arg
				245					250					255
Asn	His	Thr	Gly	Glu	Lys	Pro	Tyr	Glu	Cys	Thr	Glu	Cys	Asp	Lys
				260					265					270
Ala	Phe	Arg	Trp	Lys	Ser	Gln	Leu	Asn	Ala	His	Gln	Lys	Ile	His
				275					280					285
Thr	Gly	Glu	Lys	Ser	Tyr	Ile	Cys	Ser	Asp	Cys	Gly	Lys	Gly	Phe
				290					295					300
Ile	Lys	Lys	Ser	Arg	Leu	Ile	Asn	His	Gln	Arg	Val	His	Thr	Gly
				305					310					315

Glu	Lys	Pro	His	Gly	Cys	Ser	Leu	Cys	Gly	Lys	Ala	Phe	Ser	Lys	
				320					325					330	
Arg	Ser	Arg	Leu	Thr	Glu	His	Gln	Arg	Thr	His	Thr	Gly	Glu	Lys	
				335					340					345	
Pro	Tyr	Glu	Cys	Thr	Glu	Cys	Asp	Lys	Ala	Phe	Arg	Trp	Lys	Ser	
				350					355					360	
Gln	Leu	Asn	Ala	His	Gln	Lys	Ala	His	Thr	Gly	Glu	Lys	Ser	Tyr	
				365					370					375	
Ile	Cys	Arg	Asp	Cys	Gly	Lys	Gly	Phe	Ile	Gln	Lys	Gly	Asn	Leu	
				380					385					390	
Ile	Val	His	Gln	Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Tyr	Ile	Cys	
				395					400					405	
Asn	Glu	Cys	Gly	Lys	Gly	Phe	Ile	Gln	Lys	Gly	Asn	Leu	Leu	Ile	
				410					415					420	
His	Arg	Arg	Thr	His	Thr	Gly	Glu	Lys	Pro	Tyr	Val	Cys	Asn	Glu	
				425					430					435	
Cys	Gly	Lys	Gly	Phe	Ser	Gln	Lys	Thr	Cys	Leu	Ile	Ser	His	Gln	
				440					445					450	
Arg	Phe	His	Thr	Gly	Lys	Thr	Pro	Phe	Val	Cys	Thr	Glu	Cys	Gly	
				455					460					465	
Lys	Ser	Cys	Ser	His	Lys	Ser	Gly	Leu	Ile	Asn	His	Gln	Arg	Ile	
				470					475					480	
His	Thr	Gly	Glu	Lys	Pro	Tyr	Thr	Cys	Ser	Asp	Cys	Gly	Lys	Ala	
				485					490					495	
Phe	Arg	Asp	Lys	Ser	Cys	Leu	Asn	Arg	His	Arg	Arg	Thr	His	Thr	
				500					505					510	
Gly	Glu	Arg	Pro	Tyr	Gly	Cys	Ser	Asp	Cys	Gly	Lys	Ala	Phe	Ser	
				515					520					525	
His	Leu	Ser	Cys	Leu	Val	Tyr	His	Lys	Gly	Met	Leu	His	Ala	Arg	
				530					535					540	
Glu	Lys	Cys	Val	Gly	Ser	Val	Lys	Leu	Glu	Asn	Pro	Cys	Ser	Glu	
				545					550					555	
Ser	His	Ser	Leu	Ser	His	Thr	Arg	Asp	Leu	Ile	Gln	Asp	Lys	Asp	
				560					565					570	
Ser	Val	Asn	Met	Val	Thr	Leu	Gln	Met	Pro	Ser	Val	Ala	Ala	Gln	
				575					580					585	
Thr	Ser	Leu	Thr	Asn	Ser	Ala	Phe	Gln	Ala	Glu	Ser	Lys	Val	Ala	
				590					595					600	
Ile	Val	Ser	Gln	Pro	Val	Ala	Arg	Ser	Ser	Val	Ser	Ala	Asp	Ser	
				605					610					615	
Arg	Ile	Cys	Thr	Glu											
				620											

&lt;210&gt; 394

&lt;211&gt; 134

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1178269.2.orf3:2000SEP08

&lt;400&gt; 394

Phe	Pro	Gly	Arg	Pro	Thr	Arg	Pro	Val	Ser	Ser	Ala	Arg	Thr	Arg	
1				5					10					15	
Thr	Arg	Pro	Gly	Glu	Ala	Glu	Gly	Leu	Cys	Thr	Cys	Pro	Arg	Glu	
				20					25					30	
Glu	Ser	Arg	Glu	Ile	Arg	Ala	Gly	Gln	Ile	Val	Leu	Lys	Ala	Met	
				35					40					45	
Ala	Gln	Gly	Leu	Val	Thr	Phe	Arg	Asp	Val	Ala	Ile	Glu	Phe	Ser	
				50					55					60	
Leu	Glu	Glu	Trp	Lys	Cys	Leu	Glu	Pro	Ala	Gln	Arg	Asp	Leu	Tyr	
				65					70					75	

Arg	Glu	Val	Thr	Leu	Glu	Asn	Phe	Gly	His	Leu	Ala	Ser	Leu	Gly	
				80					85					90	
Lys	Glu	Ile	Val	Ser	His	Gly	Pro	Asp	Gly	Lys	Gln	Gly	Asn	Arg	
				95					100					105	
Asp	Tyr	Asn	His	Gly	Val	Gln	Phe	Pro	Ser	Leu	Pro	Lys	Trp	Val	
				110					115					120	
Arg	Arg	Gly	Asn	Cys	His	Arg	Arg	Arg	Tyr	Lys	Leu	Gln	Pro		
				125					130						

&lt;210&gt; 395

&lt;211&gt; 95

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:813422.1.orf1:2000SEP08

&lt;400&gt; 395

Phe	His	Leu	Met	Cys	Gly	Phe	Gln	Ser	Ile	Gln	Ile	Arg	Ala	Gly	
1				5					10					15	
Ala	Phe	Val	Ala	Leu	Ala	Pro	Glu	Pro	Ile	Gln	Phe	Leu	Phe	Leu	
				20					25					30	
Phe	Leu	Ile	Pro	Ala	Arg	Thr	Phe	Gln	Glu	Asn	Gly	Lys	Thr	Val	
				35					40					45	
Ala	Pro	Pro	Lys	Cys	Ile	Trp	Gly	Ser	Leu	Lys	Phe	Glu	Arg	Leu	
				50					55					60	
Ser	Val	Ser	Leu	Leu	Val	Gln	Ala	Pro	Trp	Pro	Phe	Leu	Gln	Phe	
				65					70					75	
Cys	Phe	Cys	His	Met	Cys	Leu	Lys	Gly	Asn	Gly	Gln	Ala	Phe	Glu	
				80					85					90	
Ile	Val	Thr	Gln	His											
				95											

&lt;210&gt; 396

&lt;211&gt; 114

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1093049.6.orf2:2000SEP08

&lt;400&gt; 396

Arg	Thr	Leu	Tyr	Arg	Asn	Val	Met	Leu	Glu	Thr	Tyr	Asn	Ser	Leu	
1				5					10					15	
Val	Ser	Leu	Gln	Glu	Leu	Val	Ser	Phe	Glu	Glu	Val	Ala	Val	His	
				20					25					30	
Phe	Thr	Trp	Glu	Glu	Trp	Gln	Asp	Leu	Asp	Asp	Ala	Gln	Arg	Thr	
				35					40					45	
Leu	Tyr	Arg	Asp	Val	Met	Leu	Glu	Thr	Tyr	Ser	Ser	Leu	Val	Ser	
				50					55					60	
Leu	Gly	His	Cys	Ile	Thr	Lys	Pro	Glu	Met	Ile	Phe	Lys	Leu	Glu	
				65					70					75	
Gln	Gly	Ala	Glu	Pro	Trp	Ile	Val	Glu	Glu	Thr	Leu	Asn	Leu	Arg	
				80					85					90	
Leu	Ser	Gly	Gly	Ser	Lys	Lys	Gln	Val	Phe	Ser	Gly	Ile	Cys	His	
				95					100					105	
Arg	Ser	Leu	Val	Glu	Leu	Gln	Glu	Val							
				110											

&lt;210&gt; 397

&lt;211&gt; 84



<212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:202192.4.orf3:2000SEP08

<400> 397  
 Pro Tyr Gly Asn Leu Met Ile Ser Leu Met Gln His Gln Ala Pro  
 1 5 10 15  
 Gly Ile Glu Arg Asp Val Gln Arg Gln Gly Pro Ala Pro Arg Gly  
 20 25 30  
 Leu Gln Ser Gln Pro Thr Ser Ala Val Val Arg Ser Ala Leu  
 35 40 45  
 Leu Val Gln His Thr Thr Gln Pro Pro Ala Val Ser His Phe Phe  
 50 55 60  
 Tyr Arg Val His His Lys Ser Pro Glu Ile Leu Val Cys Pro Arg  
 65 70 75  
 Phe His Phe Ile Thr Pro Ala Leu Arg  
 80

<210> 398  
 <211> 208  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1041854.1.orf3:2000SEP08

<400> 398  
 Gly Asn Cys Ala Ser Gln Lys Thr Pro Arg Val Val Leu Leu Ala  
 1 5 10 15  
 Leu Pro Glu Gly Arg Leu Ala Ile Pro Val Cys Leu Leu Ser Ala  
 20 25 30  
 Lys Leu Lys Ser Thr Asp Val Ser Ala Thr Met Leu Pro Ala Ala  
 35 40 45  
 Leu Leu Arg His Pro Gly Leu Arg Arg Leu Val Leu Gln Ala Arg  
 50 55 60  
 Thr Tyr Ala Gln Ala Ala Ala Ser Pro Ala Pro Ala Ala Gly Pro  
 65 70 75  
 Gly Gln Met Ser Phe Thr Phe Ala Ser Pro Thr Gln Val Phe Phe  
 80 85 90  
 Asp Gly Ala Asn Val Arg Gln Val Asp Val Pro Thr Leu Thr Gly  
 95 100 105  
 Ala Phe Gly Ile Leu Ala Ser His Val Pro Thr Leu Gln Val Leu  
 110 115 120  
 Arg Pro Gly Leu Val Val Val His Ala Glu Asp Gly Thr Thr Thr  
 125 130 135  
 Lys Tyr Phe Val Ser Ser Gly Ser Val Thr Val Asn Ala Asp Ser  
 140 145 150  
 Ser Val Gln Leu Leu Ala Glu Glu Val Val Thr Leu Asp Met Leu  
 155 160 165  
 Asp Leu Gly Ala Ala Arg Ala Asn Leu Glu Lys Ala Gln Ser Glu  
 170 175 180  
 Leu Ser Gly Ala Ala Asp Glu Ala Ala Arg Ala Glu Ile Gln Ile  
 185 190 195  
 Arg Ile Glu Ala Asn Glu Ala Leu Val Lys Ala Leu Glu  
 200 205

<210> 399  
 <211> 104  
 <212> PRT



	230		235		240
Ala Asn Glu Val Arg Lys Val Leu Asn Arg Thr Lys Glu Thr Lys					
	245		250		255
Ile Thr Gly Ala Asn Ala Ser Leu Lys Ser His Ala Pro Leu Val					
	260		265		270
His Leu Ser Ala Leu Pro Glu Gly Thr Val Val Asp Ser Arg					
	275		280		

&lt;210&gt; 401

&lt;211&gt; 297

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:400517.4.orf2:2000SEP08

&lt;400&gt; 401

Gly Ala Gly Ser Trp Leu Ser Gln Glu Gln Asp Ala Leu Ala Gln			
1 5 10 15			
Pro Gln Pro Trp Trp Lys Thr Gln Leu Phe Met Trp Glu Pro Val			
20 25 30			
Leu Phe Gly Thr Trp Asp Gly Val Phe Thr Ser Cys Met Ile Asn			
35 40 45			
Ile Phe Gly Val Val Leu Phe Leu Arg Thr Gly Trp Leu Val Gly			
50 55 60			
Asn Thr Gly Ala Leu Leu Gly Met Phe Leu Val Ser Phe Val Ile			
65 70 75			
Leu Val Ala Leu Val Thr Val Leu Ser Asp Ile Gly Val Gly Glu			
80 85 90			
Arg Ser Ser Ile Gly Ser Gly Gly Val Tyr Ser Met Ile Ser Ser			
95 100 105			
Val Leu Gly Gly Gln Thr Gly Gly Thr Ile Gly Leu Leu Tyr Val			
110 115 120			
Phe Gly Gln Val Pro Cys Ile Ser Pro Ala Leu Leu Asn Pro Ser			
125 130 135			
Arg Ile Cys Trp Ala Ser Gly Ile Ser Gly Leu Cys Glu Glu Phe			
140 145 150			
Gln Leu Arg Cys Phe Trp Pro Cys Leu Gly Ile Asn Leu Ala Gly			
155 160 165			
Val Lys Trp Ile Ile Arg Leu Gln Leu Leu Leu Leu Phe Leu Leu			
170 175 180			
Ala Val Ser Thr Leu Asp Phe Val Val Gly Ser Phe Thr His Leu			
185 190 195			
Asp Pro Glu His Gly Phe Ile Gly Tyr Ser Pro Glu Leu Leu Gln			
200 205 210			
Asn Asn Thr Leu Pro Asp Tyr Ser Pro Gly Glu Ser Phe Phe Thr			
215 220 225			
Val Phe Gly Val Phe Phe Pro Ala Ala Thr Gly Val Met Ala Gly			
230 235 240			
Phe Asn Met Gly Gly Arg Pro Gln Gly Ala Cys Arg Pro Ala Phe			
245 250 255			
Pro Trp Ala Pro Trp Gln Leu Leu Cys Ile Ser Lys Lys Leu Leu			
260 265 270			
Cys Pro Ser Leu Gly Phe Cys Thr Ser Val Phe Val Phe Leu			
275 280 285			
Leu Gly Arg His Leu His Ser Arg Gly Pro Ser Leu			
290 295			

&lt;210&gt; 402

&lt;211&gt; 148

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1078917.1.orf2:2000SEP08

&lt;400&gt; 402

```

Asp Lys Ile Val Asn Gly Glu Ala Asp Ala Met Ser Leu Asp Gly
 1          5          10          15
Gly His Ala Tyr Ile Ala Gly Gln Cys Gly Leu Val Pro Val Met
          20          25          30
Ala Glu Asn Tyr Asp Ile Ser Ser Cys Thr Asn Pro Gln Ser Asp
          35          40          45
Val Phe Pro Lys Gly Tyr Tyr Ala Val Ala Val Val Lys Ala Ser
          50          55          60
Asp Ser Ser Ile Asn Trp Asn Asn Leu Lys Gly Lys Lys Ser Cys
          65          70          75
His Thr Gly Val Asp Arg Thr Ala Gly Trp Asn Ile Pro Met Gly
          80          85          90
Leu Leu Phe Ser Arg Ile Asn His Cys Lys Phe Asp Glu Phe Phe
          95          100          105
Ser Gln Gly Cys Ala Pro Gly Tyr Lys Lys Asn Ser Thr Leu Cys
          110          115          120
Asp Leu Cys Ile Gly Pro Ala Lys Cys Ala Pro Asn Asn Arg Glu
          125          130          135
Gly Tyr Asn Gly Tyr Thr Arg Gly Phe Pro Val Pro Arg
          140          145

```

&lt;210&gt; 403

&lt;211&gt; 165

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1012560.1.orf3:2000SEP08

&lt;400&gt; 403

```

Ile Lys Leu His Ser Glu Glu Pro Cys Phe Ala Leu Asn Val Ile
 1          5          10          15
His Ile Phe Thr Arg Ser Val Val Glu Val Gly Phe Met Ile Gly
          20          25          30
Gln Tyr Leu Leu Tyr Gly Phe His Leu Glu Pro Leu Phe Lys Cys
          35          40          45
His Gly His Pro Cys Pro Asn Ile Ile Asp Cys Phe Val Ser Arg
          50          55          60
Pro Thr Glu Lys Thr Ile Phe Leu Leu Phe Met Gln Ser Ile Ala
          65          70          75
Thr Ile Ser Leu Phe Leu Asn Ile Leu Glu Ile Phe His Leu Gly
          80          85          90
Phe Lys Lys Ile Lys Arg Gly Leu Trp Gly Lys Tyr Lys Leu Lys
          95          100          105
Lys Glu His Asn Glu Phe His Ala Asn Lys Ala Lys Gln Asn Val
          110          115          120
Ala Lys Tyr Gln Ser Thr Ser Ala Asn Ser Leu Lys Arg Leu Pro
          125          130          135
Ser Ala Pro Asp Tyr Asp Leu Leu Val Glu Lys Gln Thr His Thr
          140          145          150
Ala Val Tyr Pro Ser Leu Asn Ser Ser Ser Val Phe Gln Pro Asn
          155          160          165

```

&lt;210&gt; 404

&lt;211&gt; 285

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:427997.4.orf3:2000SEP08

&lt;400&gt; 404

Arg	Thr	Gly	Gly	Gly	Gly	Ser	Gln	Ser	Pro	Ala	Pro	Asp	Gly	1	5	10	15	
Glu	Arg	Thr	Met	His	Cys	Leu	Gly	Ala	Glu	Tyr	Leu	Val	Ser	Ala	20	25	30	35
Asp	Gly	Ala	Pro	Arg	Gln	Arg	Glu	Trp	Arg	Pro	Gln	Ile	Tyr	Arg	40	45	50	55
Lys	Cys	Thr	Asp	Thr	Ala	Trp	Leu	Phe	Leu	Phe	Phe	Leu	Phe	Trp	60	65	70	75
Thr	Gly	Leu	Val	Phe	Ile	Met	Gly	Tyr	Ser	Val	Val	Ala	Gly	Ala	80	85	90	95
Ala	Gly	Arg	Leu	Leu	Phe	Gly	Tyr	Asp	Ser	Phe	Gly	Asn	Met	Cys	100	105	110	115
Gly	Lys	Lys	Asn	Ser	Pro	Val	Glu	Gly	Ala	Pro	Leu	Ser	Gly	Gln	120	125	130	135
Asp	Met	Thr	Leu	Lys	Lys	His	Val	Phe	Phe	Met	Asn	Ser	Cys	Asn	140	145	150	155
Leu	Glu	Val	Lys	Gly	Thr	Gln	Leu	Asn	Arg	Met	Ala	Leu	Cys	Val	160	165	170	175
Ser	Asn	Cys	Pro	Glu	Glu	Gln	Leu	Asp	Ser	Leu	Glu	Glu	Val	Gln	180	185	190	195
Phe	Phe	Ala	Asn	Thr	Ser	Gly	Ser	Phe	Leu	Cys	Gly	Tyr	Ser	Leu	200	205	210	215
Asn	Ser	Phe	Asn	Tyr	Thr	His	Ser	Pro	Lys	Ala	Asp	Ser	Leu	Cys	220	225	230	235
Pro	Arg	Leu	Pro	Val	Pro	Pro	Ser	Lys	Ser	Phe	Pro	Leu	Phe	Asn	240	245	250	255
Arg	Cys	Val	Pro	Gln	Thr	Pro	Glu	Cys	Tyr	Ser	Leu	Phe	Ala	Ser	260	265	270	275
Val	Leu	Ile	Asn	Asp	Val	Asp	Thr	Leu	His	Arg	Ile	Leu	Ser	Gly	280	285		

&lt;210&gt; 405

&lt;211&gt; 414

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:197899.1.orf2:2000SEP08

&lt;400&gt; 405

Ser	Leu	Ala	Ser	Leu	Arg	Gly	Thr	Cys	Thr	Thr	Gly	Ser	Ser	Gln	1	5	10	15
Lys	Arg	Thr	Thr	Asn	Trp	Pro	Trp	Leu	Asn	Ser	Met	Thr	Pro	Ile	20	25	30	35
Ala	Trp	Val	Leu	Trp	Ala	Leu	Glu	Glu	Phe	Val	Pro	Phe	Gly	Phe	40	45		

Glu	Gly	Ile	Leu	Arg	Gly	Ala	Ala	Thr	Cys	Phe	Tyr	Ala	Phe	Val	
				50					55					60	
Gly	Phe	Asp	Cys	Ile	Ala	Thr	Thr	Gly	Glu	Glu	Ala	Gln	Asn	Pro	
				65					70					75	
Gln	Arg	Ser	Ile	Pro	Met	Gly	Ile	Val	Ile	Ser	Leu	Ser	Val	Cys	
				80					85					90	
Phe	Leu	Ala	Tyr	Phe	Ala	Val	Ser	Ser	Ala	Leu	Thr	Leu	Met	Met	
				95					100					105	
Pro	Tyr	Tyr	Gln	Leu	Gln	Pro	Glu	Ser	Pro	Leu	Pro	Glu	Ala	Phe	
				110					115					120	
Leu	Tyr	Ile	Gly	Trp	Ala	Pro	Ala	Arg	Tyr	Val	Val	Ala	Val	Gly	
				125					130					135	
Ser	Leu	Cys	Ala	Leu	Ser	Thr	Ser	Leu	Leu	Gly	Ser	Met	Phe	Pro	
				140					145					150	
Met	Pro	Arg	Val	Ile	Tyr	Ala	Met	Ala	Glu	Asp	Gly	Leu	Leu	Phe	
				155					160					165	
Arg	Val	Leu	Ala	Arg	Ile	His	Thr	Gly	Thr	Arg	Thr	Pro	Ile	Ile	
				170					175					180	
Ala	Thr	Val	Val	Ser	Gly	Ile	Ile	Ala	Ala	Phe	Met	Ala	Phe	Leu	
				185					190					195	
Phe	Lys	Leu	Thr	Asp	Leu	Val	Asp	Leu	Met	Ser	Ile	Gly	Thr	Leu	
				200					205					210	
Leu	Ala	Tyr	Ser	Leu	Val	Ser	Ile	Cys	Val	Leu	Ile	Leu	Arg	Tyr	
				215					220					225	
Gln	Pro	Asp	Gln	Glu	Thr	Lys	Thr	Gly	Glu	Glu	Val	Glu	Leu	Gln	
				230					235					240	
Glu	Glu	Ala	Ile	Thr	Thr	Glu	Ser	Glu	Lys	Leu	Thr	Leu	Trp	Gly	
				245					250					255	
Leu	Phe	Phe	Pro	Leu	Asn	Ser	Ile	Pro	Thr	Pro	Leu	Ser	Gly	Gln	
				260					265					270	
Ile	Val	Tyr	Val	Cys	Ser	Ser	Leu	Leu	Ala	Val	Leu	Leu	Thr	Ala	
				275					280					285	
Leu	Cys	Leu	Val	Leu	Ala	Gln	Trp	Ser	Val	Pro	Leu	Leu	Ser	Gly	
				290					295					300	
Asp	Leu	Leu	Trp	Thr	Ala	Val	Val	Val	Leu	Leu	Leu	Leu	Leu	Ile	
				305					310					315	
Ile	Gly	Ile	Ile	Val	Val	Ile	Trp	Arg	Gln	Pro	Gln	Ser	Ser	Thr	
				320					325					330	
Pro	Leu	His	Phe	Lys	Val	Pro	Ala	Leu	Pro	Leu	Leu	Pro	Leu	Met	
				335					340					345	
Ser	Ile	Phe	Val	Asn	Ile	Tyr	Leu	Met	Met	Gln	Met	Thr	Ala	Gly	
				350					355					360	
Thr	Trp	Ala	Arg	Phe	Gly	Val	Trp	Met	Leu	Ile	Gly	Phe	Ala	Ile	
				365					370					375	
Tyr	Phe	Gly	Tyr	Gly	Ile	Gln	His	Ser	Leu	Glu	Glu	Ile	Lys	Ser	
				380					385					390	
Asn	Gln	Pro	Ser	Arg	Lys	Ser	Arg	Ala	Lys	Thr	Val	Asp	Leu	Asp	
				395					400					405	
Pro	Gly	Thr	Leu	Val	His	Ser	Val								
				410											

&lt;210&gt; 406

&lt;211&gt; 140

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:334199.1.orf3:2000SEP08

&lt;400&gt; 406

Ser	Trp	Gly	Gln	Trp	Arg	Leu	Cys	Val	Gln	Asp	Asn	Pro	Trp	Leu
1						5				10				15

```

Gly Ala Arg Arg Thr Pro Arg Gly Val Gln Arg Pro Arg Val Ala
      20                      25                      30
Arg Ala Ala Ala Ala Arg Ala Ala Pro Ala Ser Ala Ala Arg Ser
      35                      40                      45
Gly Arg Ala Arg Ala Pro Ala Arg Arg Ala Pro Gln Thr Ala Arg
      50                      55                      60
Ser Pro Arg Pro Cys Ser Arg Arg Arg Tyr Pro Arg Gly Ser Ser
      65                      70                      75
Ser Ala Ser Ser Arg Arg Arg Cys Thr His Pro Pro Glu Ala Pro
      80                      85                      90
Ala Leu Gly Pro Pro Ser Pro Ala Val Ala Arg Pro Pro Leu Pro
      95                      100                     105
Leu Asp Ala Leu Leu Arg Arg Gly Ala Ala Ser Gln Gln Pro Thr
      110                     115                     120
Gln Thr Pro Leu His Gly His Glu Pro Gln Leu Pro Ala Gln Pro
      125                     130                     135
Pro Gln Pro Ala Thr
      140

```

&lt;210&gt; 407

&lt;211&gt; 188

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:334345.1.orf2:2000SEP08

&lt;400&gt; 407

```

Arg Pro Gly Ala Val His Asn Thr Lys Arg Met Asp Ser Asp Pro
  1                      5                      10                      15
Ser Ala Val Ser Val Gly Lys Arg Ser His Gly Ala Ser Gly Arg
      20                      25                      30
Pro His Pro Ala Ser Ala Ala Val Ala Ala Arg Met Arg Pro
      35                      40                      45
Leu Gln Gly Gly Arg Glu Asp Cys Gly Arg Pro Arg His Pro Gly
      50                      55                      60
Arg Thr Leu Ala Val Ala Gly Trp Pro Val Val Asp Leu Ser Gly
      65                      70                      75
Ala Cys Met Trp Gly Leu Pro His Pro Pro Thr Leu Gly Ala His
      80                      85                      90
Ser Arg Pro Leu Leu Pro Glu Gly Asp Ser Gly Gly Pro Leu Val
      95                      100                     105
Cys Pro Ile Asn Asp Thr Trp Ile Gln Ala Gly Ile Val Ser Trp
      110                     115                     120
Gly Phe Gly Cys Ala Arg Pro Phe Arg Pro Gly Val Tyr Thr Gln
      125                     130                     135
Val Leu Ser Tyr Thr Asp Trp Ile Gln Arg Thr Leu Ala Glu Ser
      140                     145                     150
His Ser Gly Met Ser Gly Ala Arg Pro Gly Ala Pro Gly Ser His
      155                     160                     165
Ser Gly Thr Ser Arg Ser His Pro Val Leu Leu Leu Glu Leu Leu
      170                     175                     180
Thr Val Cys Leu Leu Gly Ser Leu
      185

```

&lt;210&gt; 408

&lt;211&gt; 72

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: LG:228092.1.orf3:2000SEP08

<400> 408

Ala	Leu	Gln	Ala	Val	Ala	Phe	Pro	Val	Val	Thr	Leu	Phe	Leu	Phe	
1				5						10				15	
Pro	Thr	Arg	Glu	Leu	Gly	Arg	Asn	Ala	Leu	Ala	Ala	Leu	Ser	Pro	
			20						25					30	
Gln	Ala	Gly	Ser	Arg	Cys	Ser	Lys	Arg	Pro	Arg	Pro	Leu	Arg	Ser	
			35						40					45	
Phe	Leu	Asp	Asp	Leu	Thr	Gly	Pro	Gln	Ala	Ser	Leu	Gln	Arg	Ser	
			50						55					60	
Glu	Lys	Lys	Met	Asp	Pro	Phe	His	Glu	Lys	Phe	Leu				
			65						70						

<210> 409

<211> 109

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:098580.1.orf3:2000SEP08

<400> 409

Gln	Trp	Lys	Lys	Pro	Val	Gln	Ile	Leu	Met	Val	Gly	Ser	Cys	Lys	
1				5					10					15	
Val	Thr	Ser	Val	Met	Met	Leu	Leu	Gln	Arg	Leu	Met	Trp	Glu	Lys	
			20						25					30	
Asp	Phe	Ile	Ala	Phe	Lys	Ser	Ser	Thr	Pro	His	Asn	Val	Ser	Trp	
			35						40					45	
Arg	His	Glu	Thr	Asn	Gly	Ser	Val	Phe	Ile	Ser	Gln	Ile	Ile	Tyr	
			50						55					60	
Tyr	Phe	Arg	Glu	Tyr	Ser	Trp	Ser	His	His	Leu	Glu	Glu	Ile	Phe	
			65						70					75	
Gln	Lys	Val	Gln	His	Ser	Phe	Glu	Thr	Pro	Asn	Ile	Leu	Thr	Gln	
			80						85					90	
Leu	Pro	Thr	Ile	Glu	Arg	Leu	Ser	Met	Thr	Arg	Tyr	Phe	Tyr	Leu	
			95						100					105	
Phe	Pro	Gly	Asn												

<210> 410

<211> 235

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LG:969572.1.orf3:2000SEP08

<400> 410

Gly	Met	Asn	Tyr	Lys	Ile	Ile	Tyr	Ser	Ile	Val	Gln	Thr	Asn	Cys	
1				5					10					15	
Ser	Lys	Glu	Asp	Phe	Pro	Phe	Leu	Arg	Glu	Asp	Cys	Val	Pro	Leu	
			20						25					30	
Pro	Tyr	Gly	Asp	His	Gly	Glu	Cys	Arg	Gly	His	Thr	Tyr	Val	Asp	
			35						40					45	
Ile	His	Asn	Thr	Ile	Ala	Gly	Phe	Ser	Gln	Ser	Cys	Asp	Leu	Tyr	
			50						55					60	
Pro	Gly	Asp	Asp	Leu	Phe	Ser	Leu	Leu	Pro	Lys	Lys	Cys	Phe	Gly	
			65						70					75	
Cys	Pro	Lys	Asn	Ile	Pro	Val	Asp	Ser	Pro	Glu	Leu	Lys	Glu	Ala	
			80						85					90	



```

Leu Gly His Ser Ile Ala Gln Leu Asn Ala Gln His Asn His Leu
95 100
Phe Tyr Phe Lys Ile Asp Thr Val Lys Lys Ala Thr Ser Gln Val
110 115 120
Val Ala Gly Thr Lys Tyr Val Ile Glu Phe Ile Ala Arg Glu Thr
125 130 135
Asn Cys Ser Lys Gln Thr Asn Thr Glu Leu Thr Ala Asp Cys Glu
140 145 150
Thr Lys His Leu Gly Gln Ser Leu Asn Cys Asn Ala Asn Val Tyr
155 160 165
Met Arg Pro Trp Glu Asn Lys Val Val Pro Thr Val Arg Cys Gln
170 175 180
Ala Leu Asp Met Met Ile Ser Arg Pro Pro Gly Phe Ser Pro Phe
185 190 195
Arg Leu Val Gln Val Gln Glu Thr Lys Glu Gly Thr Thr Arg Leu
200 205 210
Leu Asn Ser Cys Glu Tyr Lys Gly Arg Leu Ser Lys Ala Gly Ala
215 220 225
Gly Pro Ala Pro Asp His Gln Ala Glu Ala
230 235

```

&lt;210&gt; 411

&lt;211&gt; 210

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:196958.1.orf2:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 18, 43

&lt;223&gt; unknown or other

&lt;400&gt; 411

```

Ala Pro Gly Leu Leu Ser Arg Pro Arg Met Pro Lys Tyr Leu Gly
1 5 10 15
Gly Gly Xaa Cys Ile Pro Gly Pro Trp Ala Glu Arg Arg Val Tyr
20 25 30
Ser Leu Gly His Gln Asp Lys Ser Arg Thr His Gln Xaa Leu Arg
35 40 45
Thr Asp Arg Arg Thr Thr Glu Gly Val Thr Gly Trp Cys Glu Asp
50 55 60
Trp Cys Pro Trp Ala Arg Thr Leu Leu Ser Ser Pro Cys Trp Leu
65 70 75
Gln Thr Arg Val Gln Ala Leu Gly Ser Ala Thr Leu Thr Gln Pro
80 85 90
Ser Leu Glu Asp Arg Met Arg Gly Val Ser Cys Leu Gln Val Leu
95 100 105
Leu Leu Leu Val Leu Gly Ala Ala Gly Thr Gln Gly Arg Lys Ser
110 115 120
Ala Ala Cys Gly Gln Pro Arg Met Ser Ser Arg Ile Val Gly Gly
125 130 135
Arg Asp Gly Arg Asp Gly Glu Trp Pro Trp Gln Ala Ser Ile Gln
140 145 150
His Arg Gly Ala His Val Cys Gly Gly Ser Leu Ile Ala Pro Gln
155 160 165
Trp Val Leu Thr Ala Ala His Cys Phe Pro Ser Ala Pro Pro Arg
170 175 180
Val Ala Thr Ala Thr Arg Ser Lys Gly Ala Ala Ala Gly Leu Ala
185 190 195
His Leu Arg Arg Pro Leu Pro Arg Gly Arg Gly Arg Ala Pro Gly

```

200

205

210

<210> 412  
 <211> 80  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1087811.1.orf2:2000SEP08

<400> 412  
 Lys Glu Gly Cys Glu Leu Cys Arg Cys Cys Cys Arg Pro Gly Pro  
 1 5 10 15  
 Pro Ala Ser Ser His Gly Gln Pro His His Val Phe Pro His Cys  
 20 25 30  
 Cys Arg Trp Arg Ala Leu Gly Leu Cys Leu Leu Arg Gly Lys Gly  
 35 40 45  
 Pro Gly Lys Gln Glu Val Thr Ala His Leu Ile His Lys Thr Met  
 50 55 60  
 Leu Thr Asp Trp Ser Cys Leu Gln Thr Arg Phe Gln Arg Gln Gln  
 65 70 75  
 Lys Ile Ser Met Leu  
 80

<210> 413  
 <211> 184  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1327885.1.orf3:2000SEP08

<400> 413  
 Ala Pro Arg Asp Pro Ala Met Val Arg Ala Gly Ala Val Gly Thr  
 1 5 10 15  
 His Leu Pro Thr Ser Ser Leu Asp Ile Phe Gly Asp Leu Arg Lys  
 20 25 30  
 Met Asn Lys Arg Gln Leu Tyr Tyr Gln Val Leu Asn Phe Ala Met  
 35 40 45  
 Ile Val Ser Ser Ala Leu Met Ile Trp Lys Gly Leu Ile Val Leu  
 50 55 60  
 Thr Gly Ser Glu Ser Pro Ile Val Val Val Leu Ser Gly Ser Met  
 65 70 75  
 Glu Pro Ala Phe His Arg Gly Asp Leu Leu Phe Leu Thr Asn Phe  
 80 85 90  
 Arg Glu Asp Pro Ile Arg Ala Gly Glu Ile Val Val Phe Lys Val  
 95 100 105  
 Glu Gly Arg Asp Ile Pro Ile Val His Arg Val Ile Lys Val His  
 110 115 120  
 Glu Lys Asp Asn Gly Asp Ile Lys Phe Leu Thr Lys Gly Asp Asn  
 125 130 135  
 Asn Glu Val Asp Asp Arg Gly Leu Tyr Lys Glu Gly Gln Asn Trp  
 140 145 150  
 Leu Glu Lys Lys Asp Val Val Gly Arg Ala Arg Gly Phe Leu Pro  
 155 160 165  
 Tyr Val Gly Met Val Thr Ile Ile Met Asn Asp Tyr Pro Lys Phe  
 170 175 180  
 Lys Tyr Ala Leu

<210> 414  
 <211> 263  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:449393.1.orf3:2000SEP08

<400> 414  
 Ala Met Ser Leu Arg Val Leu Asn Pro Asn Ala Glu Val Leu Asn  
 1 5 .10 15  
 Lys Ser Ala Ala Leu His Met Asn Ile Asn Ala Ala Lys Gly Leu  
 20 25 30  
 Gln Asp Val Leu Lys Thr Asn Leu Gly Pro Lys Gly Thr Ile Lys  
 35 40 45  
 Met Leu Val Gly Gly Ala Gly Asp Leu Lys Leu Thr Lys Asp Gly  
 50 55 60  
 Asn Thr Leu Leu Lys Glu Met Gln Ile Gln Asn Pro Thr Ala Ile  
 65 70 75  
 Met Ile Ala Arg Thr Ala Val Ala Gln Asp Asp Thr Ser Gly Asp  
 80 85 90  
 Gly Thr Thr Ser Thr Val Leu Phe Ile Gly Glu Leu Met Lys Gln  
 95 100 105  
 Ser Glu Arg Cys Ile Asp Glu Gly Thr His Pro Arg Phe Leu Val  
 110 115 120  
 Asp Gly Phe Asp Val Ala Lys Arg Ala Cys Leu Asp Phe Leu Asp  
 125 130 135  
 Lys Phe Lys Thr Pro Val Val Thr Gly Glu Glu Pro Asp Arg Asp  
 140 145 150  
 Thr Leu Lys Met Val Ala Arg Thr Thr Leu Arg Thr Lys Leu Tyr  
 155 160 165  
 Glu Gly Leu Ala Asp Gln Leu Thr Asp Ile Val Val Asn Ala Val  
 170 175 180  
 Leu Cys Ile Arg Lys Pro Asp Glu Pro Ile Asp Leu Phe Met Val  
 185 190 195  
 Glu Ile Met His Met Arg His Lys Phe Asp Val Asp Thr Arg Leu  
 200 205 210  
 Val Glu Gly Leu Val Leu Asp His Gly Ser Arg His Pro Asp Met  
 215 220 225  
 Lys Arg Arg Ala Glu Asn Cys Tyr Ile Leu Thr Cys Asn Val Ser  
 230 235 240  
 Leu Glu Tyr Glu Lys Ser Glu Ile Asn Ala Gly Phe Phe Tyr Ser  
 245 250 255  
 Asn Ala Glu Gln Lys Lys Lys Lys  
 260

<210> 415  
 <211> 163  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:897616.1.orf3:2000SEP08

<400> 415  
 Gly Ile Gly Ser Gly Gly Ala Ala Arg Ser Phe Leu Val Gly Gly  
 1 5 10 15  
 Gly Val Ala Cys Ala Tyr Arg Asp Met Arg Leu Ser Trp Ile Leu  
 20 25 30  
 Thr Val Leu Ser Ile Cys Leu Ser Ala Leu Val Thr Ala Thr Gly  
 35 40 45

```

Ala Glu Gly Lys Arg Lys Leu Gln Ile Gly Val Lys Lys Arg Val
      50      55      60
Asp His Cys Pro Ile Lys Ser Arg Lys Gly Asp Val Leu His Met
      65      70      75
His Tyr Thr Gly Lys Leu Glu Asp Gly Thr Glu Phe Asp Ser Ser
      80      85      90
Leu Pro Gln Asn Gln Pro Phe Val Phe Ser Leu Gly Thr Gly Gln
      95     100     105
Val Ile Lys Gly Trp Asp Gln Gly Leu Leu Gly Met Cys Glu Gly
     110     115     120
Glu Lys Arg Lys Leu Val Ile Pro Ser Glu Leu Gly Tyr Gly Glu
     125     130     135
Arg Gly Ala Pro Pro Lys Ile Pro Gly Gly Ala Thr Leu Val Phe
     140     145     150
Glu Val Glu Leu Leu Lys Ile Glu Arg Arg Ser Glu Leu
     155     160

```

&lt;210&gt; 416

&lt;211&gt; 85

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:736860.1.orf3:2000SEP08

&lt;400&gt; 416

```

Ser Pro Ser Ala Ala Glu His His Glu Leu His Arg Arg Arg Arg
  1      5      10      15
Arg Trp Arg Gln Glu Val Leu Ala Gly Ser Gly Arg Ala Gln Pro
     20      25      30
Gly Arg Ser Gln Glu Gly Asp Pro Val Arg Gln Ala Arg Arg Arg
     35      40      45
His Arg Arg Ala Ala Arg Arg His Ala Gly Asp His Gly Phe Pro
     50      55      60
Pro Gln Pro Arg Pro His Leu Arg Arg His Arg Arg Gly Gly Thr
     65      70      75
Pro His Arg Leu Arg Phe Lys Ser Thr Lys
     80      85

```

&lt;210&gt; 417

&lt;211&gt; 239

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:027066.6.orf3:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 37, 62, 64

&lt;223&gt; unknown or other

&lt;400&gt; 417

```

Val Gly Gly Arg Leu Cys Glu Gly Gly Gln Glu Val Trp Gly Pro
  1      5      10      15
Gly Ala Pro Ala Gly Leu Leu Pro Arg Ala Ala Leu Gln Arg Val
     20      25      30
His Leu Ser Cys Gly Gly Xaa Glu Glu Leu Ser Gly Gly Arg Phe
     35      40      45
Gly Ala Ser Asn Leu Glu Pro Ala Thr Trp Val Asp Pro Gly Arg
     50      55      60

```

```

Arg Xaa Leu Xaa Ser Gly Cys Gly Ala Pro Arg Glu Leu Gly Pro
      65      70      75
Pro Arg Ser Ser His Ser Ala Asp Glu Gly Leu Arg Gly Arg Arg
      80      85      90
Gln Gly Thr Gly Val Ile Gln Thr Ala Trp Ala Thr His Trp Leu
      95     100     105
Pro Cys His Thr Ala Lys Gly Phe Ser Pro Ser Leu Gly Ser Gly
     110     115     120
Ser Val Ala Pro Ser Gly Pro Leu Leu Ala Ser Pro Leu Ala Tyr
     125     130     135
Gly Gly Gly Arg Leu Leu Ser Ser Trp Ser Arg Pro Ser Gln Pro
     140     145     150
Ser Phe Cys Cys Leu Pro Thr Arg Ser Leu Pro Thr Arg Ser Leu
     155     160     165
Pro Ala Thr Pro Ser Ser Leu Gln Thr Ala Ile Val Thr Gly Asp
     170     175     180
Gly Ala Gly Trp Arg Gly Pro Gly Arg Ser Asn Pro Gln Ala Leu
     185     190     195
Cys Pro Pro Gly Leu Pro Arg Thr Glu Thr Thr Ala Pro Ala Pro
     200     205     210
Val Leu Ala Ser Ser Ser Pro His Thr Ser Ala Ser Thr Gly Pro
     215     220     225
Glu Ser Gly Ser Thr Trp Thr Leu Leu His Arg Cys Met Leu
     230     235

```

&lt;210&gt; 418

&lt;211&gt; 142

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1074263.1.orf3:2000SEP08

&lt;400&gt; 418

```

Gly Thr Met Ala Ser Pro Leu Arg Ser Leu Met Leu Leu Leu Ala
  1      5      10      15
Val Leu Ala Val Ala Trp Ala Gly Thr Ser Arg Pro Pro Pro Arg
  20      25      30
Leu Leu Gly Ala Pro Gln Glu Ala Asp Ala Ser Glu Glu Gly Val
  35      40      45
Gln Arg Ala Leu Asp Phe Ala Val Ser Glu Tyr Asn Lys Gly Ser
  50      55      60
Asn Asp Ala Tyr His Ser Arg Ala Ile Gln Val Val Arg Ala Arg
  65      70      75
Lys Gln Leu Val Ala Gly Ile Asn Tyr Tyr Leu Asp Val Glu Met
  80      85      90
Gly Arg Thr Thr Cys Thr Lys Ser Gln Thr Asn Leu Thr Asn Cys
  95     100     105
Pro Phe His Asp Gln Pro His Leu Met Arg Lys Ala Leu Cys Ser
  110     115     120
Phe Gln Ile Tyr Ser Val Pro Trp Lys Gly Thr His Thr Leu Thr
  125     130     135
Lys Ser Ser Cys Lys Asn Ala
  140

```

&lt;210&gt; 419

&lt;211&gt; 191

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: LI:334345.1.orf1:2000SEP08

<400> 419

Arg	Gly	Arg	Asp	Glu	Asp	Ser	Asp	Pro	Ser	Ala	Val	Ser	Val	Gly
1				5					10					15
Lys	Arg	Ser	His	Gly	Ala	Ser	Gly	Arg	Pro	His	Pro	Ala	Ser	Ala
				20					25					30
Ala	Val	Ala	Ala	Ala	Arg	Gly	Pro	Leu	Gly	Cys	Ala	Ala	Phe	Arg
				35					40					45
Met	Arg	Pro	Leu	Gln	Gly	Gly	Arg	Glu	Asp	Cys	Gly	Arg	Pro	Arg
				50					55					60
His	Pro	Gly	Arg	Thr	Leu	Asp	Val	Ala	Gly	Trp	Pro	Val	Val	Asp
				65					70					75
Leu	Ser	Gly	Ala	Cys	Met	Trp	Gly	Leu	Pro	His	Pro	Pro	Thr	Leu
				80					85					90
Gly	Ala	His	Ser	Arg	Pro	Leu	Leu	Pro	Glu	Gly	Asp	Ser	Gly	Gly
				95					100					105
Pro	Leu	Val	Cys	Pro	Ile	Asn	Asp	Thr	Trp	Ile	Gln	Ala	Gly	Ile
				110					115					120
Val	Ser	Trp	Gly	Phe	Gly	Cys	Ala	Arg	Pro	Phe	Arg	Pro	Gly	Val
				125					130					135
Tyr	Thr	Gln	Val	Leu	Ser	Tyr	Thr	Asp	Trp	Ile	Gln	Arg	Thr	Leu
				140					145					150
Ala	Glu	Ser	His	Ser	Gly	Met	Ser	Gly	Ala	Arg	Pro	Gly	Ala	Pro
				155					160					165
Gly	Ser	His	Ser	Gly	Thr	Ser	Arg	Ser	His	Pro	Val	Leu	Leu	Leu
				170					175					180
Glu	Leu	Leu	Thr	Val	Cys	Leu	Leu	Gly	Ser	Leu				
				185					190					

<210> 420

<211> 88

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1093914.1.orf2:2000SEP08

<400> 420

Lys	Lys	Leu	Val	Cys	Ile	Met	Leu	Phe	Tyr	Pro	Arg	Tyr	Ala	Ala
1				5					10					15
Gln	Gly	Ile	Thr	Asn	Leu	Met	Arg	Ile	Met	Thr	His	Phe	Lys	Pro
				20					25					30
Pro	Met	Phe	Thr	Val	Phe	Lys	Ile	Lys	Leu	Arg	Thr	Gly	Pro	Phe
				35					40					45
Leu	Gly	Asp	Thr	Ser	Lys	Val	Ile	Ala	Arg	Thr	Glu	Glu	Arg	Lys
				50					55					60
Val	Pro	Lys	Asn	Val	Thr	Leu	Lys	Phe	Asp	Ala	Cys	Ala	Ala	Ile
				65					70					75
Asn	Ser	Lys	Gln	His	Gly	Ile	Gly	Cys	Gly	Ser	Leu	Asp		
				80					85					

<210> 421

<211> 152

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1188168.1.orf2:2000SEP08

<400> 421

```

Arg Ala Val Ala Gly Ser Gly Gly Gln Ser Gly Gln Cys Leu Gly
 1          5          10          15
Ala Ala Gly Ser Thr Gly Leu Val Gly Arg Leu Leu His Leu Lys
          20          25          30
Asp Pro Val Ser Asp Gln Leu Leu Ala Cys Ser Ala Pro Val Ala
          35          40          45
Gly Ala Pro Gly Glu Leu Gly Glu Phe His Ala Phe Ile Arg Leu
          50          55          60
Gly Pro Val Gly Ser Ser Arg Ala Gln Ala Gly Gln Cys Ile Leu
          65          70          75
Leu Leu Ala Asp Pro Glu Pro Leu Gly Ser Val Gly Cys Trp Gln
          80          85          90
Gly Leu Trp Arg Glu Gly Gly Glu Gly Trp Ser Gln Val Asp Ala
          95          100          105
Ala Val Ala Ser Glu Leu Pro Val Pro Ala Pro Gly Arg Gly Cys
          110          115          120
Ser Gly Trp Arg Arg Arg Ser Ser Ser Gly Ser Leu Thr Ser Ser
          125          130          135
Pro Leu Ala Thr Arg Ser Arg Arg Pro Ala Thr Cys Arg Ser Leu
          140          145          150
Tyr Thr

```

<210> 422  
 <211> 91  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1065168.1.orf1:2000SEP08

```

<400> 422
Pro Gln Met Leu Gln Thr Asn Arg His Lys Leu Val Ile Gly Cys
 1          5          10          15
Ser Gly Phe His Gly Asp Cys Leu Thr Leu Thr Lys Ile Ile Glu
          20          25          30
Ala Arg Leu Lys Met Tyr Lys His Ser Asn Asn Lys Ala Met Thr
          35          40          45
Thr Gly Ala Ile Ala Ala Met Leu Ser Thr Ile Leu Tyr Ser Arg
          50          55          60
Arg Phe Phe Pro Tyr Tyr Val Tyr Asn Ile Gln Phe Arg Lys Asp
          65          70          75
Ile Asp Glu Glu Lys Gly Lys Thr Val Tyr Ser Phe Glu Pro
          80          85          90
Lys

```

<210> 423  
 <211> 80  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1180418.1.orf2:2000SEP08

```

<400> 423
Lys Glu Gly Cys Glu Leu Cys Arg Cys Cys Cys Arg Pro Gly Pro
 1          5          10          15
Pro Ala Ser Ser His Gly Gln Pro His His Val Phe Pro His Cys
          20          25          30
Cys Arg Trp Arg Ala Leu Gly Leu Cys Leu Leu Arg Gly Lys Gly

```

	35							40					45
Pro Gly Lys Gln Glu Val Thr Ala His Leu Ile His Lys Thr Met													
	50							55					60
Leu Thr Asp Trp Ser Cys Leu Gln Thr Arg Phe Gln Arg Gln Gln													
	65							70					75
Lys Ile Ser Met Leu													
	80												

&lt;210&gt; 424

&lt;211&gt; 362

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:232648.1.orf1:2000SEP08

&lt;400&gt; 424

Gly Cys Arg Cys Thr Val Gly Leu Arg Thr Arg Leu Gly Gly Phe													
1	5							10					15
Ala Pro Glu Pro Gly Trp Gly Gly Ala Gly Thr Ser Val Cys Trp													
	20							25					30
Ala Gly Gly Ala Pro Arg Ala His Ser Thr Arg Thr His Ala Trp													
	35							40					45
Leu Gly Arg Pro Ala Met Glu Pro Gly Ser Val Glu Asn Leu Ser													
	50							55					60
Ile Val Tyr Arg Ser Arg Asp Phe Leu Val Val Asn Lys His Trp													
	65							70					75
Asp Val Arg Ile Asp Ser Lys Ala Trp Arg Glu Thr Leu Thr Leu													
	80							85					90
Gln Lys Gln Leu Arg Tyr Arg Phe Pro Glu Leu Ala Asp Pro Asp													
	95							100					105
Thr Cys Tyr Gly Phe Arg Phe Cys His Gln Leu Asp Phe Ser Thr													
	110							115					120
Ser Gly Ala Leu Cys Val Ala Leu Asn Lys Ala Ala Ala Gly Ser													
	125							130					135
Ala Tyr Arg Cys Phe Lys Glu Arg Arg Val Thr Lys Ala Tyr Leu													
	140							145					150
Ala Leu Leu Arg Gly His Ile Gln Glu Ser Arg Val Thr Ile Ser													
	155							160					165
His Ala Ile Gly Arg Asn Ser Thr Glu Gly Arg Ala His Thr Met													
	170							175					180
Cys Ile Glu Gly Ser Gln Gly Cys Glu Asn Pro Lys Pro Ser Leu													
	185							190					195
Thr Asp Leu Val Val Leu Glu His Gly Leu Tyr Ala Gly Asp Pro													
	200							205					210
Val Ser Lys Val Leu Leu Lys Pro Leu Thr Gly Arg Thr His Gln													
	215							220					225
Leu Arg Val His Cys Ser Ala Leu Gly His Pro Val Val Gly Asp													
	230							235					240
Leu Thr Tyr Gly Glu Val Ser Gly Arg Glu Asp Arg Pro Phe Arg													
	245							250					255
Met Met Leu His Ala Phe Tyr Leu Arg Ile Pro Thr Asp Thr Glu													
	260							265					270
Cys Val Glu Val Cys Thr Pro Asp Pro Phe Leu Pro Ser Leu Asp													
	275							280					285
Ala Cys Trp Ser Pro His Thr Leu Leu Gln Ser Leu Asp Gln Leu													
	290							295					300
Val Gln Ala Leu Arg Ala Thr Pro Asp Pro Asp Pro Glu Asp Arg													
	305							310					315
Gly Pro Arg Pro Gly Ser Pro Ser Ala Leu Leu Pro Gly Pro Gly													
	320							325					330
Arg Pro Pro Pro Pro Pro Thr Lys Pro Pro Glu Thr Glu Ala Gln													



	335		340		345
Arg Gly Pro Cys	Leu Gln Trp Leu Ser	Glu Trp Thr Leu Glu	Pro		
	350		355		360
Asp Ser					

<210> 425  
 <211> 169  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1078420.1.orf2:2000SEP08

<400> 425

Asn Ser Gly Glu Thr Met Glu Met Met Leu Asp Lys Lys Gln Ile		
1	5	10
Arg Ala Ile Phe Leu Phe Lys Phe Lys Met Gly His Lys Ala Ala		15
	20	25
Glu Thr Thr His Asn Ile Ser Asn Thr Phe Gly Pro Gly Thr Ala		30
	35	40
Asn Glu Arg Thr Val Gln Trp Trp Phe Lys Lys Phe His Lys Gly		45
	50	55
Glu Glu Ser Leu Glu Asp Glu Glu His Ser Gly Gln Pro Ser Glu		60
	65	70
Val Asp Ser Asp Gln Leu Arg Ala Ile Ile Glu Ala Asp Pro Leu		75
	80	85
Ile Thr Thr Pro Glu Val Ala Glu Glu Leu Asn Ile Asp His Ser		90
	95	100
Thr Val Val Arg His Leu Lys Gln Ile Gly Lys Val Lys Lys Leu		105
	110	115
Gly Lys Trp Val Pro Cys Glu Leu Ser Glu Lys Gln Lys Asn Cys		120
	125	130
Arg Phe Glu Val Ser Ser Ser Leu Ile Leu Arg Asn Asn Ser Glu		135
	140	145
Pro Phe Leu Asp Gln Ile Val Met Cys His Glu Lys Trp Ile Leu		150
	155	160
Tyr Asp Ser Gln		165

<210> 426  
 <211> 95  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1397599.1.orf2:2000SEP08

<400> 426

Leu Leu Arg Tyr Phe Ala Glu Lys Gly Arg Asp Val Phe Ser Ser		
1	5	10
Asn Ala Ile Lys Leu Val Asn Val Thr Arg Val His Val Lys Arg		15
	20	25
Leu Pro Asn Arg Leu Cys Val Ser Asn Asn Ala Val Tyr Phe Gly		30
	35	40
Ala Glu Asp Pro Gly Gln Glu Asp Ser Phe Arg Arg Arg Val Pro		45
	50	55
Cys Pro Cys Pro His Ser Val Arg Arg Ser Thr Tyr Asp Leu Gly		60
	65	70
Ser Ser Asp Gln Pro Ala Gln Gly Thr Ser His Glu Phe Gln Ile		75
	80	85
		90

**<220>**

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1101065.1.orf1:2000SEP08

&lt;400&gt; 429

Lys	Tyr	Gly	Gly	Lys	Leu	Val	Cys	Ile	Val	Cys	Ile	Leu	Ile	Val
1				5					10					15
Pro	Pro	Thr	Gly	Cys	Ser	Ser	Ile	Pro	Leu	Pro	Leu	Leu	Glu	Pro
				20					25					30
Pro	Cys	Ser	Leu	Arg	His	Asn	Asn	Ile	Glu	Ile	Gly	Pro	Ile	Asn
				35					40					45
Asn	Pro	Thr	Met	Thr	Ser	Lys	Cys	Ser	Asn	Glu	Arg	Lys	Ser	Ser
				50					55					60
Thr														

&lt;210&gt; 430

&lt;211&gt; 101

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:475629.1.orf2:2000SEP08

&lt;400&gt; 430

Thr	Phe	Arg	Gly	Gly	Gly	Glu	Val	Pro	Ala	Ala	Ala	Ala	Met	Ala
1				5					10					15
Ser	Thr	Lys	Val	Gln	Arg	Ile	Met	Thr	Gln	Pro	Ile	Asn	Leu	Ile
				20					25					30
Phe	Arg	Phe	Leu	Gln	Ser	Lys	Ala	Arg	Ile	Gln	Ile	Trp	Leu	Phe
				35					40					45
Glu	Gln	Lys	Asp	Leu	Arg	Ile	Glu	Gly	Arg	Ile	Ile	Gly	Phe	Asp
				50					55					60
Glu	Tyr	Met	Asn	Leu	Val	Leu	Glu	Asp	Ala	Glu	Glu	Ile	Asn	Val
				65					70					75
Lys	Lys	Asn	Thr	Arg	Lys	Ser	Leu	Gly	Arg	Ile	Leu	Leu	Lys	Gly
				80					85					90
Asp	Asn	Ile	Thr	Leu	Met	Met	Asn	Ser	Gly	Lys				
				95					100					

&lt;210&gt; 431

&lt;211&gt; 274

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:348991.1.orf1:2000SEP08

&lt;400&gt; 431

Gly	Arg	Gly	Arg	Gly	Arg	Gly	Gly	Gly	Glu	Gly	Lys	Asp	Arg	Gly
1				5					10					15
Gly	Gly	Gly	Gln	Arg	Arg	Gly	Gly	Val	Ala	Lys	Ser	Lys	Ser	Arg
				20					25					30
Arg	Arg	Lys	Gly	Ala	Met	Val	Val	Ser	Val	Glu	Pro	His	Arg	His
				35					40					45
Glu	Gly	Val	Phe	Ile	Tyr	Arg	Gly	Ala	Glu	Asp	Ala	Leu	Val	Thr
				50					55					60
Leu	Asn	Met	Val	Pro	Gly	Gln	Ser	Val	Tyr	Gly	Glu	Arg	Arg	Val
				65					70					75
Thr	Val	Thr	Glu	Gly	Gly	Val	Lys	Gln	Glu	Tyr	Arg	Thr	Trp	Asn
				80					85					90
Pro	Phe	Arg	Ser	Lys	Leu	Ala	Ala	Ala	Ile	Leu	Gly	Gly	Val	Asp

	95		100		105
Gln Ile His Ile	Lys Pro Lys Ser Lys	Val Leu Tyr Leu Gly	Ala		
	110		115		120
Ala Ser Gly Thr	Thr Val Ser His Val	Ser Asp Ile Ile Gly	Pro		
	125		130		135
Asp Gly Leu Val	Tyr Ala Val Asp Phe	Ser His Arg Ala Gly	Arg		
	140		145		150
Asp Leu Val Asn	Val Ala Lys Lys Arg	Thr Asn Ile Ile Pro	Val		
	155		160		165
Leu Glu Asp Ala	Arg His Pro Leu Lys	Tyr Arg Met Leu Ile	Gly		
	170		175		180
Met Val Asp Val	Ile Phe Ala Asp Val	Ala Gln Pro Asp Gln	Ser		
	185		190		195
Arg Ile Val Ala	Leu Asn Ala His Thr	Phe Leu Arg Asn Gly	Gly		
	200		205		210
His Phe Leu Ile	Ser Ile Lys Ala Asn	Cys Ile Asp Ser Thr	Ala		
	215		220		225
Ser Ala Glu Ala	Val Phe Ala Ser Glu	Val Arg Lys Leu Gln	Gln		
	230		235		240
Glu Asn Leu Lys	Pro Gln Glu Gln Leu	Thr Leu Glu Pro Tyr	Glu		
	245		250		255
Arg Asp His Ala	Val Val Val Gly Val	Tyr Arg Pro Leu Pro	Lys		
	260		265		270
Ser Ser Ser Lys					

&lt;210&gt; 432

&lt;211&gt; 110

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:475629.1.orf3:2000SEP08

&lt;400&gt; 432

Leu Gly Ser Gly	Leu Ala Ala Pro Pro	Arg Ala His Ser	Gln Ala
1	5	10	15
Ser Arg Arg Ser	Glu Glu Asp Gly	Glu Val Pro Ala	Ala Ala
	20	25	30
Met Ala Ser Thr	Lys Val Gln Arg	Ile Met Thr	Gln Pro Ile Asn
	35	40	45
Leu Ile Phe Arg	Phe Leu Gln Ser	Lys Ala Arg	Ile Gln Ile Trp
	50	55	60
Leu Phe Glu Gln	Lys Asp Leu Arg	Ile Glu Gly	Arg Ile Ile Gly
	65	70	75
Phe Asp Glu Tyr	Met Asn Leu Val	Leu Glu Asp	Ala Glu Glu Ile
	80	85	90
Asn Ser Gln Glu	Ile Thr Leu Gly	Asn His Trp	Ala Gly Tyr Ser
	95	100	105
Ser Lys Val Thr	Ile		
	110		

&lt;210&gt; 433

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:261331.1.orf1:2000SEP08

&lt;400&gt; 433

Ala Ile Phe Val Ser Lys Phe Glu Met Gly His Lys Ala Val Glu  
 1 5 10 15  
 Thr Thr Arg Asn Ile His Asn Ala Phe Gly Ala Gly Thr Ala Asn  
 20 25 30  
 Glu Leu Thr Val Gln Trp Trp Phe Lys Lys Phe Cys Lys Gly Asp  
 35 40 45  
 Lys Ser Leu Glu Asp Glu Glu His Gly Gly Gln Pro Ser Glu Val  
 50 55 60  
 Asp Asn Asp Gln Leu Lys Ala Ile Val Lys Ala Asp Pro Leu Thr  
 65 70 75  
 Thr Thr Arg Glu Val Ala Lys Glu Leu Asn Ile Asn His Ser Val  
 80 85 90  
 Val Ile Gln His Leu Lys Gln Met Glu Gly Lys Lys Ala Trp  
 95 100

&lt;210&gt; 434

&lt;211&gt; 69

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:815686.1.orf3:2000SEP08

&lt;400&gt; 434

Gly Ser Trp Glu Phe Asn His Val Asn Val Val Phe Phe Ile Leu  
 1 5 10 15  
 Leu Thr Lys Glu Gln Trp Val Pro Phe Thr Lys Phe Leu Arg Ser  
 20 25 30  
 Gly Asn Lys Lys Lys Glu Glu Pro Asn Lys Asp Cys Asn Met Ser  
 35 40 45  
 Ala Tyr Gly Phe Pro Ile Glu Ala Leu Thr Lys Leu Pro Leu Phe  
 50 55 60  
 Asp Glu Arg Asn Lys Gln Lys Pro Trp  
 65

&lt;210&gt; 435

&lt;211&gt; 74

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1167327.2.orf3:2000SEP08

&lt;400&gt; 435

Gly Asp Pro Pro Thr Thr Ser Gly Pro Gln Thr Asn Gln Pro Lys  
 1 5 10 15  
 Glu His Leu Met Asn Phe Lys Ser Asp Ser Gln Leu Tyr Glu Asp  
 20 25 30  
 Thr Leu Ala Gly Arg Ser Val Leu Ile Lys Asn Leu Thr Pro Gln  
 35 40 45  
 Thr Leu Gln Pro Arg Trp Thr Gly Pro Tyr Leu Val Ile Tyr Ser  
 50 55 60  
 Thr Pro Thr Ala Val Arg Leu Gln Asp Pro Pro His Trp Val  
 65 70

&lt;210&gt; 436

&lt;211&gt; 187

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:758009.3.orf2:2000SEP08

&lt;400&gt; 436

```

Pro Phe Trp Ser Asn Thr Cys Gln Lys Ala Arg Val Gly Gly Ser
  1          5          10          15
Val Cys Ser Met Gly Gly Lys Ala Trp Pro Leu Pro Thr Ala Ser
          20          25          30
Lys Ser Ser Arg Arg Ser Leu Gln Thr Gly Arg Ala Pro Pro Phe
          35          40          45
Leu Arg Gly Leu Glu Glu Ala Pro Ala Ser Ser Glu Glu Thr Tyr
          50          55          60
Gln Val Pro Thr Leu Pro Arg Pro Pro Thr Pro Gly Pro Val Tyr
          65          70          75
Glu Gln Met Arg Ser Trp Ala Glu Gly Pro Gln Pro Pro Thr Ala
          80          85          90
Gln Val Tyr Glu Phe Pro Asp Pro Pro Thr Ser Ala Arg Ile Ile
          95          100          105
Cys Glu Lys Thr Leu Ser Phe Pro Lys Gln Ala Ile Leu Thr Leu
          110          115          120
Pro Arg Pro Val Arg Ala Ser Leu Pro Thr Leu Pro Ser Gln Val
          125          130          135
Tyr Asp Val Pro Thr Gln His Arg Gly Pro Val Val Leu Lys Glu
          140          145          150
Pro Glu Lys Gln Gln Leu Tyr Asp Ile Pro Ala Ser Pro Lys Lys
          155          160          165
Ala Gly Leu His Pro Pro Asp Ser Gln Ala Ser Gly Gln Gly Val
          170          175          180
Pro Leu Ile Ser Val Thr Thr
          185

```

&lt;210&gt; 437

&lt;211&gt; 187

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:331593.1.orf3:2000SEP08

&lt;400&gt; 437

```

Asn Pro Arg Arg Gly Leu Gly Ser Gly Arg Arg Asp Ala Met Thr
  1          5          10          15
Ala Glu Phe Leu Ser Leu Leu Cys Leu Gly Leu Cys Leu Gly Tyr
          20          25          30
Glu Asp Glu Lys Lys Asn Glu Lys Pro Pro Lys Pro Ser Leu His
          35          40          45
Ala Trp Pro Ser Ser Val Val Glu Ala Glu Ser Asn Val Thr Leu
          50          55          60
Lys Cys Gln Ala His Ser Gln Asn Val Thr Phe Val Leu Arg Lys
          65          70          75
Val Asn Asp Ser Gly Tyr Lys Gln Glu Gln Ser Ser Ala Glu Asn
          80          85          90
Glu Ala Glu Phe Pro Phe Thr Asp Leu Lys Pro Lys Asp Ala Gly
          95          100          105
Arg Tyr Phe Cys Ala Tyr Lys Thr Thr Ala Ser His Glu Trp Ser
          110          115          120
Glu Ser Ser Glu His Leu Gln Leu Val Val Thr Asp Lys His Asp
          125          130          135
Glu Leu Glu Ala Pro Ser Met Lys Thr Asp Thr Arg Thr Ile Phe
          140          145          150
Val Ala Ile Phe Ser Cys Ile Ser Ile Leu Leu Leu Phe Leu Ser
          155          160          165

```

Val Phe Ile Ile Tyr Arg Cys Ser Gln His Ser Glu Leu Arg Glu  
 170 175 180  
 Arg Lys Gly Arg Glu Gly Glu  
 185

<210> 438  
 <211> 158  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1094174.1.orf2:2000SEP08

<400> 438  
 Arg Lys Lys Lys Thr Cys Leu Tyr Ala Ser Ile Pro Arg Gly Glu  
 1 5 10 15  
 Leu Pro Leu Trp Arg Arg Ala Pro Ser Leu Ser Pro Ser Ser Leu  
 20 25 30  
 Gly Glu Ala Asn Arg Arg Arg Trp Gly Pro Cys Ser Arg Ser Pro  
 35 40 45  
 Arg Glu Pro Thr Gly Thr Gln Ile Leu Pro Ile Arg Arg Gly Trp  
 50 55 60  
 Ala Val Met Gly Ala Pro Lys Pro Ser Phe Leu Leu Leu Leu Arg  
 65 70 75  
 Gly Pro Trp Arg Leu Asp Ala Arg Pro Trp Ala Gly Leu Pro Leu  
 80 85 90  
 Ser Met Arg Tyr Phe Tyr Thr Asp Gln Cys Pro Gly Pro Gly Arg  
 95 100 105  
 Arg Gly Ser Pro Cys Phe Ile Cys Leu Cys Gly Tyr Val Gly Arg  
 110 115 120  
 Gln Arg Met Phe Val Leu Val Arg Ser Ala Asp Arg Arg Glu Pro  
 125 130 135  
 Glu Cys Met Asp Ala Gly Gly Arg Arg Ala Ile Asp Ala Gly Gly  
 140 145 150  
 Ala Val Arg Ser Ser Gly Pro Gly  
 155

<210> 439  
 <211> 175  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:814362.1.orf3:2000SEP08

<400> 439  
 Ala Gly Arg Gly Ser Lys Met Val Leu Gln Thr Gln Val Phe Ile  
 1 5 10 15  
 Ser Leu Leu Leu Trp Ile Ser Val Leu Thr Thr Gly Ala Tyr Gly  
 20 25 30  
 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu  
 35 40 45  
 Gly Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Ser  
 50 55 60  
 Tyr Ser Ser Asp Asn Arg Asn Phe Leu Ala Trp Tyr Gln Gln Lys  
 65 70 75  
 Arg Gly Gln Pro Pro Lys Val Leu Ile Asn Glu Ala Ser Asn Arg  
 80 85 90  
 Glu Ser Gly Val Pro Glu Arg Phe Ser Gly Ser Gly Ser Gly Thr  
 95 100 105  
 Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Ala Glu Asp Val Ala

	110		115		120
Val Tyr Tyr Cys	Gln Gln Tyr Phe Ser	Leu Pro Leu Thr Phe	Gly		
	125		130		135
Gly Gly Thr Lys	Val Glu Ile Lys Arg Thr	Val Ala Ala Pro	Ser		
	140		145		150
Val Phe Ile Phe	Pro Pro Ser Asp Glu	Gln Leu Lys Ser Gly	Thr		
	155		160		165
Ala Ser Val Val	Cys Leu Leu Asn Asn	Phe			
	170		175		

<210> 440  
 <211> 87  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:219542.1.orf3:2000SEP08

<400> 440	
Arg Gln Leu Gln Ser Phe Met Met Phe Pro Leu Ala Arg Asn Ala	
1 5 10 15	
Leu Ser Ser Leu Lys Ile Gln Ser Ile Leu Gln Ser Met Ala Arg	
20 25 30	
His Ser His Val Lys His Ser Pro Asp Phe His Asp Lys Tyr Gly	
35 40 45	
Asn Ala Val Leu Ala Ser Gly Thr Ala Phe Cys Val Ala Thr Trp	
50 55 60	
Val Phe Thr Ala Thr Gln Ile Gly Ile Glu Trp Asn Leu Ser Pro	
65 70 75	
Val Gly Arg Val Thr Pro Lys Glu Trp Lys His Gln	
80 85	

<210> 441  
 <211> 119  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:726197.1.orf3:2000SEP08

<400> 441	
Ser Pro Glu Ser Arg Ala Met Ala Leu Ala Lys Ala Lys Glu Ile	
1 5 10 15	
Val Ala Ser Ala Pro Leu Val Val Phe Ser Lys Thr Ser Cys Pro	
20 25 30	
Phe Cys Val Arg Val Lys Gln Leu Phe Glu Lys Leu Gly Ala Ser	
35 40 45	
Tyr Lys Ala Ile Glu Leu Asp Lys Glu Ser Asp Gly Ala Glu Leu	
50 55 60	
Gln Asn Ala Leu Lys Glu Trp Thr Gly Gln Arg Thr Val Pro Asn	
65 70 75	
Val Phe Ile Asn Gly Lys His Ile Gly Gly Cys Asp Asp Thr Met	
80 85 90	
Ala Leu Asn Asn Asp Gly Lys Leu Val Pro Leu Leu Thr Glu Ala	
95 100 105	
Gly Ala Ile Ala Gly Ser Ala Ser Lys Thr Thr Ile Thr Ala	
110 115	

<210> 442  
 <211> 226  
 <212> PRT



&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1075314.1.orf1:2000SEP08

&lt;400&gt; 442

```

Pro Glu Leu Lys Met Ala Ala Leu Leu Leu Arg His Ile Gly Arg
 1          5          10          15
His Cys Leu Arg Ala His Leu Ser Ser Gln Leu Cys Ile Arg Asn
          20          25          30
Ala Ala Pro Leu Gly Thr Thr Ala Lys Glu Glu Met Ala Arg Phe
          35          40          45
Trp Asn Lys Asn Thr Ser Ser Asn Arg Pro Val Ser Pro His Leu
          50          55          60
Thr Ile Tyr Arg Trp Ser Leu Pro Met Ala Met Ser Val Cys His
          65          70          75
Arg Gly Ser Gly Ile Ala Met Ser Gly Gly Val Ser Leu Phe Gly
          80          85          90
Leu Ser Ala Leu Leu Leu Pro Gly Asn Phe Glu Ser Tyr Leu Met
          95          100          105
Leu Val Lys Ser Leu Cys Leu Gly Pro Ala Leu Ile His Ala Ala
          110          115          120
Lys Phe Val Leu Val Phe Pro Leu Met Tyr His Ser Leu Asn Gly
          125          130          135
Val Arg His Leu Met Trp Asp Leu Gly Lys Gly Leu Ser Ile Ser
          140          145          150
Gln Val Gln Leu Ser Gly Val Ala Gly Leu Gly Pro Gly Ser Ala
          155          160          165
Val Leu Cys Arg Thr Gly Ser His Met Lys Ser Trp Asp Ser His
          170          175          180
Ile Arg Pro Val His His His Thr Asp Leu Tyr Ser Cys Leu Ser
          185          190          195
Leu Leu Pro Pro Ala Thr Lys Val Leu Leu Ile Cys Leu Asp Val
          200          205          210
Met Cys Phe Arg Ser Pro Gly Gly Thr Val Glu Lys Leu Thr Glu
          215          220          225
Leu

```

&lt;210&gt; 443

&lt;211&gt; 152

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:437883.1.orf2:2000SEP08

&lt;400&gt; 443

```

Gly Glu Pro Arg Thr Leu Pro Cys Ser Pro Cys Ala Leu Arg Ala
 1          5          10          15
Arg Arg Pro Ala Ser Pro Leu Ser Gly Glu Asp Met Gln Gly Ile
          20          25          30
Cys Ile Arg Gln Gln Cys Lys Met Val Leu Glu Glu Pro Tyr Leu
          35          40          45
Cys Ile Glu Ile Leu Leu Arg Ile Thr Gln Ile Thr Pro Phe Asp
          50          55          60
Phe Thr Pro Glu Asn Tyr Glu Arg Tyr Arg Gly Asn Ser Leu Glu
          65          70          75
Thr Thr Arg Lys Gly Thr Glu Gln Gln Leu Trp Leu Pro Val Leu
          80          85          90
Asp Leu Ala Gln Arg Gln Asn Gly Trp Leu Pro Ile Ser Ala Met

```

	95		100		105
Asn Lys Val Ala	Glu Val Leu Gln Val	Pro Pro Met Arg Val	Tyr		
	110		115		120
Glu Val Ala Thr	Phe Tyr Thr Met Tyr	Asn Arg Lys Pro Val	Gly		
	125		130		135
Ile Val Pro His	Ser Gly Leu His Tyr	Tyr Thr Leu His Ala	Ser		
	140		145		150
Arg Leu					

<210> 444  
 <211> 172  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:336265.1.orf3:2000SEP08

<400> 444	
Arg Glu Asn Gln Glu Arg Lys Glu Ser Gln Ala Arg Arg Gly Ser	
1 5 10 15	
Leu Gly Arg Pro Gly Glu Pro Gly Phe Lys Gly Glu Arg Gly Asp	
20 25 30	
Pro Gly Ile Lys Gly Asp Lys Gly Pro Pro Gly Gly Lys Gly Gln	
35 40 45	
Pro Gly Asp Pro Gly Ile Pro Gly His Lys Gly His Thr Gly Leu	
50 55 60	
Met Gly Pro Gln Gly Leu Pro Gly Glu Asn Gly Pro Val Gly Pro	
65 70 75	
Pro Gly Pro Pro Gly Gln Pro Gly Phe Pro Gly Leu Arg Gly Glu	
80 85 90	
Ser Pro Ser Met Glu Thr Leu Arg Arg Leu Ile Gln Glu Glu Leu	
95 100 105	
Gly Lys Gln Leu Glu Thr Arg Leu Ala Tyr Leu Leu Ala Gln Met	
110 115 120	
Pro Pro Ala Tyr Met Lys Ser Ser Gln Gly Arg Pro Gly Pro Pro	
125 130 135	
Gly Pro Pro Gly Lys Asp Gly Leu Pro Gly Arg Ala Gly Pro Met	
140 145 150	
Gly Gly Ala Arg Ser Ser Trp Ala Gly Gly Ser Gly Arg Thr Leu	
155 160 165	
Trp Thr His Arg Ser Gln Arg	
170	

<210> 445  
 <211> 169  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:407788.2.orf3:2000SEP08

<400> 445	
Leu Thr Pro Thr His His Leu Asn Ile Arg Asp Gly Val Ser Leu	
1 5 10 15	
Phe Leu Lys Ala Gln Leu Pro Val Leu Leu Gln Ser Gly Arg Ile	
20 25 30	
Arg Asn Cys Asp His Cys Leu Ser Gln His Gly Ser Pro Gly Ile	
35 40 45	
Pro Gly Pro Pro Gly Pro Ile Gly Pro Glu Gly Pro Arg Gly Leu	
50 55 60	

```

Pro Gly Leu Pro Gly Arg Asp Gly Val Pro Gly Leu Val Gly Val
      65      70      75
Pro Gly Arg Pro Gly Val Arg Gly Leu Lys Gly Leu Pro Gly Arg
      80      85      90
Asn Gly Glu Lys Gly Ser Gln Gly Phe Gly Tyr Pro Gly Glu Gln
      95     100     105
Gly Pro Pro Gly Pro Pro Gly Pro Glu Gly Pro Pro Gly Ile Ser
     110     115     120
Lys Glu Gly Pro Pro Gly Asp Pro Gly Leu Pro Gly Lys Asp Gly
     125     130     135
Asp His Gly Lys Pro Gly Ile Gln Gly Gln Thr Gly Pro Pro Gly
     140     145     150
Ile Cys Asp Pro Ser Leu Cys Leu Ser Val Ile Ala Arg Arg Asp
     155     160     165
Pro Phe Arg Lys

```

&lt;210&gt; 446

&lt;211&gt; 220

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1326925.1.orf2:2000SEP08

&lt;400&gt; 446

```

Val Lys Leu Gly Val Leu Cys Glu Val Val Pro Val Val Val Ile
  1      5      10      15
Glu Glu Val Lys Leu Gly Ile Leu Trp Asp Val Ala Ser Met Val
     20     25     30
Val Ile Lys Glu Val Lys Leu Gly Val Leu Trp Asp Ala Val Ser
     35     40     45
Val Val Val Ile Lys Glu Val Arg Leu Gly Val Leu Cys Glu Val
     50     55     60
Val Leu Val Leu Ala Ile Ala Lys Val Asn Leu Gly Val Leu Cys
     65     70     75
Lys Val Val Leu Val Val Ala Ile Ala Lys Val Asn Leu Gly Gly
     80     85     90
Leu Val Glu Val Val Leu Glu Val Ala Ile Glu Glu Glu Lys Pro
     95    100    105
Gly Val Leu Cys Glu Trp Val Ser Val Val Val Leu Glu Glu Val
    110    115    120
Lys Leu Gly Gly Leu Cys Glu Val Val Ser Val Val Ala Ile Glu
    125    130    135
Glu Val Lys Leu Gly Val Leu Cys Glu Val Val Leu Val Leu Val
    140    145    150
Ile Lys Ala Val Lys Leu Gly Gly Leu Leu Glu Val Val Ser Val
    155    160    165
Val Val Met Glu Glu Val Lys Leu Gly Val Leu Cys Glu Gly Leu
    170    175    180
Ser Val Val Val Ile Glu Glu Val Lys Leu Gly Val Leu Cys Glu
    185    190    195
Val Val Phe Leu Val Val Ile Glu Glu Val Lys Leu Asp Val Leu
    200    205    210
Cys Glu Val Val Ser Val Val Met Ile Glu
    215    220

```

&lt;210&gt; 447

&lt;211&gt; 194

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:332655.2.orf1:2000SEP08

&lt;400&gt; 447

Arg	Thr	Ser	Met	Gly	Ile	Leu	Tyr	Ser	Glu	Pro	Ile	Cys	Gln	Ala	
1				5					10					15	
Ala	Tyr	Gln	Asn	Asp	Phe	Gly	Gln	Val	Trp	Arg	Trp	Val	Lys	Glu	
			20						25					30	
Asp	Ser	Ser	Tyr	Ala	Asn	Val	Gln	Asp	Gly	Phe	Asn	Gly	Asp	Thr	
			35						40					45	
Pro	Leu	Ile	Cys	Ala	Cys	Arg	Arg	Gly	His	Val	Arg	Ile	Val	Ser	
			50						55					60	
Phe	Leu	Leu	Arg	Arg	Asn	Ala	Asn	Val	Asn	Leu	Lys	Asn	Gln	Lys	
			65						70					75	
Glu	Arg	Thr	Cys	Leu	His	Tyr	Ala	Val	Lys	Lys	Lys	Phe	Thr	Phe	
			80						85					90	
Ile	Asp	Tyr	Leu	Leu	Ile	Ile	Leu	Leu	Met	Pro	Val	Leu	Leu	Ile	
			95						100					105	
Gly	Tyr	Phe	Leu	Met	Val	Ser	Lys	Thr	Lys	Gln	Asn	Glu	Ala	Leu	
			110						115					120	
Val	Arg	Met	Leu	Leu	Asp	Ala	Gly	Val	Glu	Val	Asn	Ala	Thr	Asp	
			125						130					135	
Cys	Tyr	Gly	Cys	Thr	Ala	Leu	His	Tyr	Ala	Cys	Glu	Met	Lys	Asn	
			140						145					150	
Gln	Ser	Leu	Ile	Pro	Leu	Leu	Leu	Glu	Ala	Arg	Ala	Asp	Pro	Thr	
			155						160					165	
Ile	Lys	Asn	Lys	His	Gly	Glu	Ser	Ser	Leu	Asp	Ile	Ala	Arg	Arg	
			170						175					180	
Leu	Lys	Phe	Ser	Gln	Ile	Glu	Leu	Met	Leu	Arg	Lys	Ala	Leu		
			185						190						

&lt;210&gt; 448

&lt;211&gt; 106

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1184621.4.orf2:2000SEP08

&lt;400&gt; 448

Asn	Tyr	Arg	Thr	Glu	Gln	Val	Asn	Ile	Ser	Ile	Val	Lys	Phe	Leu	
1				5					10					15	
Glu	Val	Lys	Leu	Leu	Ala	Lys	Lys	Tyr	Cys	Thr	Tyr	Phe	Asn	Val	
			20						25					30	
Arg	Met	Phe	Phe	Leu	Cys	Lys	Leu	Lys	Lys	Ile	Lys	Asn	Ile	Ile	
			35						40					45	
Asn	Thr	Asn	Ser	Pro	Phe	Lys	His	Lys	Gly	Trp	Gly	Phe	Leu	Phe	
			50						55					60	
Leu	Lys	Ala	Gln	Leu	Pro	Val	Leu	Leu	Gln	Ser	Gly	Arg	Ile	Arg	
			65						70					75	
Asn	Val	Ile	Ile	Cys	Pro	Ser	Gln	His	Gly	Ser	Pro	Gly	Ile	Pro	
			80						85					90	
Gly	Pro	Arg	Gly	Pro	Ile	Gly	Pro	Glu	Gly	Pro	Arg	Gly	Leu	Leu	
			95						100					105	
Val															

&lt;210&gt; 449

&lt;211&gt; 203

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2051386.1.orf2:2000SEP08

&lt;400&gt; 449

```

Ser Thr Leu Asn Leu Pro Glu Ser Trp Ala Ala Val Gly Cys Ser
 1          5          10          15
Gly Arg Ala Val Ala Pro Ala Val Glu Ala Val Ala Pro Ala Val
 20          25          30
Gly Ala Val Ala Pro Ala Val Glu Ala Val Ala Leu Ala Val Gly
 35          40          45
Ala Val Ala Pro Ala Ala Val Cys Pro Ser Ala Ala Ala Ser Pro
 50          55          60
Cys Ala Ala Val Cys Gln Pro Val Pro Ala Pro Ala Val Ala Pro
 65          70          75
Val Gly Ala Pro Arg Gly Thr Val Ala Leu Val Gly Ala Pro Lys
 80          85          90
Gly Ala Val Val Pro Val Gly Ala Pro Arg Gly Ala Val Val Leu
 95          100         105
Met Ala Ala Pro Ser Pro Ala Ala Ala Ser Pro Ala Ala Ala Pro
110         115         120
Gln Ala Val Gly His Pro Ala Ala Ser Pro Ala Ala Val Ser Pro
125         130         135
Ala Ala Ala Ser Pro Ala Ala Val Ser Pro Thr Ala Ala Ser Pro
140         145         150
Ala Ala Val Ser Pro Val Ala Val Pro Gln Ala Val Gly His Pro
155         160         165
Ala Gly Ser Pro Val Ala Ala Ile Pro Gly Ala Pro Ser Leu Ala
170         175         180
Ala Val Phe Pro Cys Ala Gly Ser Val Arg Ser Glu Ala Leu Thr
185         190         195
Ala Asp Cys Arg Trp Pro Asp Trp
200

```

&lt;210&gt; 450

&lt;211&gt; 160

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:362757.1.orf1:2000SEP08

&lt;400&gt; 450

```

Arg Ile Pro Pro Pro Ser Thr Val Pro Val Pro Val Pro Phe Arg
 1          5          10          15
Arg Ala Ser Ser Pro Glu Met Ala Asn Ala Arg Ser Gly Val Ala
 20          25          30
Val Asn Asp Glu Cys Met Leu Lys Phe Gly Glu Leu Gln Ser Lys
 35          40          45
Arg Leu His Arg Phe Leu Thr Phe Lys Met Asp Asp Lys Phe Lys
 50          55          60
Glu Ile Val Val Asp Gln Val Gly Asp Arg Ala Thr Ser Tyr Glu
 65          70          75
Asp Phe Thr Asn Ser Leu Pro Glu Asn Asp Cys Arg Tyr Ala Ile
 80          85          90
Tyr Asp Phe Asp Phe Val Thr Ala Glu Asp Val Gln Lys Ser Arg
 95          100         105
Ile Phe Tyr Ile Leu Trp Ser Pro Ser Ser Ala Lys Val Lys Ser
110         115         120
Lys Met Leu Tyr Ala Ser Ser Asn Gln Lys Phe Lys Ser Gly Leu
125         130         135
Asn Gly Ile Gln Val Glu Leu Gln Ala Thr Asp Ala Ser Glu Ile

```

	140		145	150
Ser	Leu	Asp	Glu	Ile
	155	Lys	Asp	Arg
		Ala	Arg	

<210> 451  
 <211> 223  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:406770.1.orf3:2000SEP08

<400> 451  
 Val Thr Ser Val Pro Ala Leu Leu Gln His Ser Arg Asp Gln Gln  
 1 5 10 15  
 Ala Pro Gln Pro Arg Asp Trp Val Cys Trp Ser Leu His Thr Val  
 20 25 30  
 Val Gly Arg Pro Ser Gly Gln Thr Ala Val Gly Pro Gln Ser Lys  
 35 40 45  
 Lys Ala Gly Glu Gly Arg Ala Gln Leu Pro Thr Pro Arg Thr Ala  
 50 55 60  
 Arg Asp Ser Pro Gly Ala Ala Gln Gly Ser Gly Pro His Trp Pro  
 65 70 75  
 Pro Ala Ala Glu Ala Asn Ala Gly His His His His Gln Leu Glu  
 80 85 90  
 Phe Arg His Gln Cys Ser Pro Ser Ile Cys Leu Gln Val Val Ala  
 95 100 105  
 Val Gly Gly Lys Gln Ser Thr Ser Gln Cys Lys Thr Lys Ile Pro  
 110 115 120  
 Ser Glu Ala Pro Glu Ala Lys Gln Gly Arg Gly Met Arg Gly Ala  
 125 130 135  
 Ala Gly His Gly Pro Pro Glu Ala Ser Gly Leu Ser Phe Ile Ala  
 140 145 150  
 His Met Asp Arg Arg Ser Cys Gly Ala Glu Asn Phe Ser Arg Asn  
 155 160 165  
 Arg Cys Arg Thr Ala Arg Asp Arg Ser Ala Gly Arg Ala Arg Cys  
 170 175 180  
 Gln Asp Ser Ala Glu Gly Lys Pro Leu Thr Leu Gln Ser Ser Gln  
 185 190 195  
 Thr Tyr Ala Leu Val Trp Thr Val Phe Pro Glu Gly Met Gly Lys  
 200 205 210  
 Gly Leu Val Leu Ala Glu Ser Val Tyr His Asp His Phe  
 215 220

<210> 452  
 <211> 267  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1094640.1.orf3:2000SEP08

<220>  
 <221> unsure  
 <222> 52, 242, 250, 257-259  
 <223> unknown or other

<400> 452  
 Arg Gly Ile Phe Pro Glu Val Glu Pro Pro Glu Ala Ile Pro Glu  
 1 5 10 15  
 Ile Pro Glu His Pro Pro Thr Glu Glu Phe Glu Val Phe Lys Glu

	20		25		30
Val Ile Pro Glu Gly	Glu Thr Pro Ile	Val Lys Arg Arg Lys	Thr		
35		40		45	
Pro Ser Pro Thr Ala	Pro Xaa Ala Met Lys	Glu Val Val Pro Glu			
50		55		60	
Met Lys Ile Phe Glu	Asp Val Pro Glu Glu	Pro Glu Thr Pro Arg			
65		70		75	
Met Lys Thr Pro Glu	Ala Pro Gln Glu Ile	Ile Pro Ala Lys Thr			
80		85		90	
Val Pro Ser Lys Lys	Arg Glu Pro Pro Ser	Val Lys Val Pro Glu			
95		100		105	
Ala Leu Gln Glu Ile	Val Pro Glu Lys Lys	Thr Leu Val Val Pro			
110		115		120	
Leu Arg Lys Pro Glu	Val Leu Pro Asp Glu	Val Pro Glu Ala Leu			
125		130		135	
Arg Glu Val Val Pro	Glu Lys Lys Val His	Pro Pro Gln Arg Ala			
140		145		150	
Glu Val Val Pro Val	Lys Val His Glu Ala	Pro Lys Glu Ile Ile			
155		160		165	
Pro Glu Lys Lys Val	Ser Val Val Pro Pro	Lys Lys Pro Glu Val			
170		175		180	
Pro Pro Val Lys Val	Pro Glu Ala Ser Lys	Glu Val Ile Arg Glu			
185		190		195	
Glu Lys Val Pro Leu	Ala Pro Pro Lys Glu	Pro Glu Val Pro Pro			
200		205		210	
Val Lys Val Pro Glu	Pro Pro Lys Glu Val	Val Pro Glu Lys Lys			
215		220		225	
Ala Pro Val Ala Pro	Pro Lys Glu Pro Glu	Val Pro Pro Val Lys			
230		235		240	
Val Xaa Glu Ala Pro	Lys Glu Val Val Xaa	Glu Lys Lys Val Pro			
245		250		255	
Val Xaa Xaa Xaa Lys	Lys Pro Glu Val Pro	Pro Thr			
260		265			

&lt;210&gt; 453

&lt;211&gt; 525

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:001929.1.orf1:2000SEP08

&lt;400&gt; 453

Leu Ser Val Ser Ala	Met Ser Leu Ser Pro	Cys Arg Ala Gln Arg		
1	5	10		15
Gly Phe Ser Ala Arg	Ser Ala Cys Ser Ala	Arg Ser Arg Gly Arg		
20		25		30
Ser Arg Gly Gly Phe	Ser Ser Arg Gly Gly	Phe Ser Ser Arg Ser		
35		40		45
Leu Asn Ser Phe Gly	Gly Cys Leu Glu Gly	Ser Arg Gly Ser Thr		
50		55		60
Trp Gly Ser Gly Gly	Arg Leu Gly Val Arg	Phe Gly Glu Trp Ser		
65		70		75
Gly Gly Pro Gly Leu	Ser Leu Cys Pro Pro	Gly Gly Ile Gln Glu		
80		85		90
Val Thr Ile Asn Gln	Asn Pro Leu Thr Pro	Leu Lys Ile Glu Ile		
95		100		105
Asp Pro Gln Phe Gln	Val Val Arg Thr Gln	Glu Thr Gln Glu Ile		
110		115		120
Arg Thr Leu Asn Asn	Gln Phe Ala Ser Phe	Ile Asp Lys Val Arg		
125		130		135
Phe Leu Glu Gln Gln	Asn Lys Val Leu Glu	Thr Lys Trp His Leu		

140	145	150
Leu Gln Gln Gln Gly	Leu Ser Gly Ser Gln Gln Gly Leu Glu	Pro
155	160	165
Val Phe Glu Ala Cys	Leu Asp Gln Leu Arg Lys Gln Leu Glu	Gln
170	175	180
Leu Gln Gly Glu Arg	Gly Ala Leu Asp Ala Glu Leu Lys Ala	Cys
185	190	195
Arg Asp Gln Glu Glu	Tyr Lys Ser Lys Tyr Glu Glu Glu	Ala
200	205	210
His Arg Arg Ala Thr	Leu Glu Asn Asp Phe Val Val Leu Lys	Lys
215	220	225
Asp Val Asp Gly Val	Phe Leu Ser Lys Met Glu Leu Glu Gly	Lys
230	235	240
Leu Glu Ala Leu Arg	Glu Tyr Leu Tyr Phe Leu Lys His Leu	Asn
245	250	255
Glu Glu Glu Leu Gly	Gln Leu Gln Thr Gln Ala Ser Asp Thr	Ser
260	265	270
Val Val Leu Ser Met	Asp Asn Asn Arg Tyr Leu Asp Phe Ser	Ser
275	280	285
Ile Ile Thr Glu Val	Arg Ala Arg Tyr Glu Glu Ile Ala Arg	Ser
290	295	300
Ser Lys Ala Glu Ala	Glu Ala Leu Tyr Gln Thr Lys Tyr Gln	Glu
305	310	315
Leu Gln Val Ser Ala	Gln Leu His Gly Asp Arg Met Gln Glu Thr	
320	325	330
Lys Val Gln Ile Ser	Gln Leu His Gln Glu Ile Gln Arg Leu	Gln
335	340	345
Ser Gln Thr Glu Asn	Leu Lys Lys Gln Asn Ala Ser Leu Gln	Ala
350	355	360
Ala Ile Thr Asp Ala	Glu Gln Arg Gly Glu Leu Ala Leu Lys	Asp
365	370	375
Ala Gln Ala Lys Val	Asp Glu Leu Glu Ala Ala Leu Arg Met	Ala
380	385	390
Lys Gln Asn Leu Ala	Arg Leu Leu Cys Glu Tyr Gln Glu Leu	Thr
395	400	405
Ser Thr Lys Leu Ser	Leu Asp Val Glu Ile Ala Thr Tyr Arg	Arg
410	415	420
Leu Leu Glu Gly Glu	Glu Cys Arg Met Ser Gly Glu Cys Thr	Ser
425	430	435
Gln Val Thr Ile Ser	Ser Val Gly Gly Ser Ala Val Met Ser	Gly
440	445	450
Gly Val Gly Gly Gly	Leu Gly Ser Thr Cys Gly Leu Gly Ser	Gly
455	460	465
Lys Gly Ser Pro Gly	Ser Cys Cys Thr Ser Ile Val Thr Gly	Gly
470	475	480
Ser Asn Ile Ile Leu	Gly Ser Gly Lys Asp Pro Val Leu Asp	Ser
485	490	495
Cys Ser Val Ser Gly	Ser Ser Ala Gly Ser Ser Cys His Thr	Ile
500	505	510
Leu Lys Lys Thr Val	Glu Ser Ser Leu Lys Thr Ser Ile Thr	Tyr
515	520	525

&lt;210&gt; 454

&lt;211&gt; 142

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:401322.1.orf3:2000SEP08

&lt;400&gt; 454



```

Gly Pro Ala Ser Asn Arg Ala Leu Gly Phe Val Val Phe Ala Gly
 1          5          10          15
Asp Pro Tyr Cys Val Trp Ser Gln Gly Thr Phe Gln Leu Gln Leu
          20          25          30
Leu Asp Gln Pro Cys Val Ser Ala Ser Pro Ser Met Trp Ala Arg
          35          40          45
Pro Glu Cys Arg Trp Ala Met Pro Ala Gly Ser Cys Thr Ala Trp
          50          55          60
Ser Thr Thr Ser Ser Pro Val Ala Pro Cys Pro Ala Thr Arg Pro
          65          70          75
Trp Gly Ala Ser Asp Asn Ser Phe Asn Thr Phe Phe Arg Glu Thr
          80          85          90
Gln Pro Gly Arg His Val Ser Trp Ala Val Cys Gly Pro Gly Ala
          95          100          105
Cys Cys His Arg Leu Ala Ser Thr Thr Gln Ser Pro Thr Val Val
          110          115          120
Pro Gly Ala Cys Cys Ser Gln Gly Ala Ala Gly Ser Leu Arg Ala
          125          130          135
Lys Gln Tyr His Ser His His
          140

```

&lt;210&gt; 455

&lt;211&gt; 154

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:208748.1.orf1:2000SEP08

&lt;400&gt; 455

```

Gln Lys Ile Arg Asp Gln Leu Pro Ser Trp Ile Asp Gln Glu Arg
 1          5          10          15
Ser Trp Ala Val Ala Thr Leu Lys Ile Ala Leu Pro Thr Ser Leu
          20          25          30
Tyr Arg His Leu Cys Phe Glu Asp Ala Ala Leu Trp Arg Thr Tyr
          35          40          45
Tyr Asn Asn Ser Met Cys Glu Gln Glu Phe Pro Ser Ile Leu Ala
          50          55          60
Lys Lys Val Ser Leu Phe Gln Gln Ile Leu Val Val Gln Val Leu
          65          70          75
Arg Pro Asp Arg Leu Gln Ser Ala Met Ala Leu Phe Ala Cys Lys
          80          85          90
Thr Leu Gly Leu Lys Glu Val Ser Pro Leu Pro Leu Asn Leu Lys
          95          100          105
Arg Leu Tyr Lys Glu Thr Leu Glu Ile Glu Pro Ile Leu Ile Ile
          110          115          120
Ile Ser Pro Gly Ala Asp Pro Ser Gln Glu Leu Gln Glu Leu Ala
          125          130          135
Asn Ala Glu Lys Lys Arg Arg Val Leu Ser Pro Gly Cys His Gly
          140          145          150
Ser Arg Ser Ser

```

&lt;210&gt; 456

&lt;211&gt; 596

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:407242.1.orf2:2000SEP08

&lt;400&gt; 456

Leu	Ala	Ala	Ser	Glu	Arg	Leu	Phe	Gln	Arg	Trp	Ser	Glu	Gly	Pro
1				5					10					15
Glu	Ala	Ser	Pro	Leu	Pro	Leu	Ala	Thr	Ser	Pro	Arg	Ala	Ser	Ala
				20					25					30
Ala	Ser	Arg	Ala	Pro	Ala	Gly	Ala	Met	Ala	Leu	Lys	Gly	Gln	Glu
				35					40					45
Asp	Tyr	Ile	Tyr	Leu	Phe	Lys	Asp	Ser	Thr	His	Pro	Val	Asp	Phe
				50					55					60
Leu	Asp	Ala	Phe	Arg	Thr	Phe	Tyr	Leu	Asp	Gly	Leu	Phe	Thr	Asp
				65					70					75
Ile	Thr	Leu	Gln	Cys	Pro	Ser	Gly	Ile	Ile	Phe	His	Cys	His	Arg
				80					85					90
Ala	Val	Leu	Ala	Ala	Cys	Ser	Asn	Tyr	Phe	Lys	Ala	Met	Phe	Thr
				95					100					105
Ala	Asp	Met	Lys	Glu	Lys	Phe	Lys	Asn	Lys	Ile	Lys	Leu	Ser	Gly
				110					115					120
Ile	His	His	Asp	Ile	Leu	Glu	Gly	Leu	Val	Asn	Tyr	Ala	Tyr	Thr
				125					130					135
Ser	Gln	Ile	Glu	Ile	Thr	Lys	Arg	Asn	Val	Gln	Ser	Leu	Leu	Glu
				140					145					150
Ala	Ala	Asp	Leu	Leu	Gln	Phe	Leu	Ser	Val	Lys	Lys	Ala	Cys	Glu
				155					160					165
Arg	Phe	Leu	Val	Arg	His	Leu	Asp	Ile	Asp	Asn	Cys	Ile	Gly	Met
				170					175					180
His	Ser	Phe	Ala	Glu	Phe	His	Val	Cys	Pro	Glu	Leu	Glu	Lys	Glu
				185					190					195
Ser	Arg	Arg	Ile	Leu	Cys	Ser	Lys	Phe	Lys	Glu	Val	Trp	Gln	Gln
				200					205					210
Glu	Glu	Phe	Leu	Glu	Ile	Ser	Leu	Glu	Lys	Phe	Leu	Phe	Ile	Leu
				215					220					225
Ser	Arg	Lys	Asn	Leu	Ser	Val	Trp	Lys	Glu	Glu	Ala	Ile	Ile	Glu
				230					235					240
Pro	Val	Ile	Lys	Trp	Thr	Ala	His	Asp	Val	Glu	Asn	Arg	Ile	Glu
				245					250					255
Cys	Leu	Tyr	Asn	Leu	Leu	Ser	Tyr	Ile	Asn	Ile	Asp	Ile	Asp	Pro
				260					265					270
Val	Tyr	Leu	Lys	Thr	Ala	Leu	Gly	Leu	Gln	Arg	Ser	Cys	Leu	Leu
				275					280					285
Thr	Glu	Asn	Lys	Ile	Arg	Ser	Leu	Ile	Tyr	Asn	Ala	Leu	Asn	Pro
				290					295					300
Met	His	Lys	Glu	Ile	Ser	Gln	Arg	Ser	Thr	Ala	Thr	Met	Tyr	Ile
				305					310					315
Ile	Gly	Gly	Tyr	Tyr	Trp	His	Pro	Leu	Ser	Glu	Val	His	Ile	Trp
				320					325					330
Asp	Pro	Leu	Thr	Asn	Val	Trp	Ile	Gln	Gly	Ala	Glu	Ile	Pro	Asp
				335					340					345
Tyr	Thr	Arg	Glu	Ser	Tyr	Gly	Val	Thr	Cys	Leu	Gly	Pro	Asn	Ile
				350					355					360
Tyr	Val	Thr	Gly	Gly	Tyr	Arg	Thr	Asp	Asn	Ile	Glu	Ala	Leu	Asp
				365					370					375
Thr	Val	Trp	Ile	Tyr	Asn	Ser	Glu	Ser	Asp	Glu	Trp	Thr	Glu	Gly
				380					385					390
Leu	Pro	Met	Leu	Asn	Ala	Arg	Tyr	Tyr	His	Cys	Ala	Val	Thr	Leu
				395					400					405
Gly	Gly	Cys	Val	Tyr	Ala	Leu	Gly	Gly	Tyr	Arg	Lys	Gly	Ala	Pro
				410					415					420
Ala	Glu	Glu	Ala	Glu	Phe	Tyr	Asp	Pro	Leu	Lys	Glu	Lys	Trp	Ile
				425					430					435
Pro	Ile	Ala	Asn	Met	Ile	Lys	Gly	Val	Gly	Asn	Ala	Thr	Ala	Cys
				440					445					450
Val	Leu	His	Asp	Val	Ile	Tyr	Val	Ile	Gly	Gly	His	Cys	Gly	Tyr
				455					460					465

Arg Gly Ser Cys Thr Tyr Asp Lys Val Gln Ser Tyr Asn Ser Asp	470	475	480
Ile Asn Glu Trp Ser Leu Ile Thr Ser Ser Pro His Pro Glu Tyr	485	490	495
Gly Leu Cys Ser Val Pro Phe Glu Asn Lys Leu Tyr Leu Val Gly	500	505	510
Gly Gln Thr Thr Ile Thr Glu Cys Tyr Asp Pro Glu Gln Asn Glu	515	520	525
Trp Arg Glu Ile Ala Pro Met Met Glu Arg Arg Met Glu Cys Gly	530	535	540
Ala Val Ile Met Asn Gly Cys Ile Tyr Val Thr Gly Gly Tyr Ser	545	550	555
Tyr Ser Lys Gly Thr Tyr Leu Gln Ser Ile Glu Lys Tyr Asp Pro	560	565	570
Asp Leu Asn Lys Trp Glu Ile Val Gly Asn Leu Pro Ser Ala Met	575	580	585
Arg Ser His Gly Cys Val Cys Val Tyr Asn Val	590	595	

&lt;210&gt; 457

&lt;211&gt; 762

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:403409.1.orf3:2000SEP08

&lt;400&gt; 457

Gly Arg Gly Arg Arg Lys Pro Asn Glu Phe Leu Gly Gly Cys Arg	1	5	10	15
Met Gly Asp Ser Lys Val Lys Val Ala Val Arg Ile Arg Pro Met	20	25	30	
Asn Arg Arg Glu Thr Asp Leu His Thr Lys Cys Val Val Asp Val	35	40	45	
Asp Ala Asn Lys Val Ile Leu Asn Pro Val Asn Thr Asn Leu Ser	50	55	60	
Lys Gly Asp Ala Arg Gly Gln Pro Lys Val Phe Ala Tyr Asp His	65	70	75	
Cys Phe Trp Ser Met Asp Glu Ser Val Lys Glu Lys Tyr Ala Gly	80	85	90	
Gln Asp Ile Val Phe Lys Cys Leu Gly Glu Asn Ile Leu Gln Asn	95	100	105	
Ala Phe Asp Gly Tyr Asn Ala Cys Ile Phe Ala Tyr Gly Gln Thr	110	115	120	
Gly Ser Gly Lys Ser Tyr Thr Met Met Gly Thr Ala Asp Gln Pro	125	130	135	
Gly Leu Ile Pro Arg Leu Cys Ser Gly Leu Phe Glu Arg Thr Gln	140	145	150	
Lys Glu Glu Asn Glu Glu Gln Ser Phe Lys Val Glu Val Ser Tyr	155	160	165	
Met Glu Ile Tyr Asn Glu Lys Val Arg Asp Leu Leu Asp Pro Lys	170	175	180	
Gly Ser Arg Gln Thr Leu Lys Val Arg Glu His Ser Val Leu Gly	185	190	195	
Pro Tyr Val Asp Gly Leu Ser Lys Leu Ala Val Thr Ser Tyr Lys	200	205	210	
Asp Ile Glu Ser Leu Met Ser Glu Gly Asn Lys Ser Arg Thr Val	215	220	225	
Ala Ala Thr Asn Met Asn Glu Glu Ser Ser Arg Ser His Ala Val	230	235	240	
Phe Lys Ile Thr Leu Thr His Thr Leu Tyr Asp Val Lys Ser Gly	245	250	255	

Thr	Ser	Gly	Glu	Lys	Val	Gly	Lys	Leu	Ser	Leu	Val	Asp	Leu	Ala
				260					265					270
Gly	Ser	Glu	Arg	Ala	Thr	Lys	Thr	Gly	Ala	Ala	Gly	Asp	Arg	Leu
				275					280					285
Lys	Glu	Gly	Ser	Asn	Ile	Asn	Lys	Ser	Leu	Thr	Thr	Leu	Asp	Leu
				290					295					300
Val	Ile	Ser	Ala	Leu	Ala	Asp	Gln	Ser	Ala	Gly	Lys	Asn	Lys	Asn
				305					310					315
Lys	Phe	Val	Pro	Tyr	Arg	Asp	Ser	Val	Leu	Thr	Trp	Leu	Leu	Lys
				320					325					330
Asp	Ser	Leu	Gly	Gly	Asn	Ser	Lys	Thr	Ala	Met	Val	Ala	Thr	Val
				335					340					345
Ser	Pro	Ala	Ala	Asp	Asn	Tyr	Asp	Glu	Thr	Leu	Ser	Thr	Leu	Arg
				350					355					360
Tyr	Ala	Asp	Arg	Ala	Lys	His	Ile	Val	Asn	His	Ala	Val	Val	Asn
				365					370					375
Glu	Asp	Pro	Asn	Ala	Arg	Ile	Ile	Arg	Asp	Leu	Arg	Glu	Glu	Val
				380					385					390
Glu	Lys	Leu	Arg	Glu	Gln	Leu	Thr	Lys	Ala	Glu	Ala	Met	Lys	Ser
				395					400					405
Pro	Glu	Leu	Lys	Asp	Arg	Leu	Glu	Glu	Ser	Glu	Lys	Leu	Ile	Gln
				410					415					420
Glu	Met	Thr	Val	Thr	Trp	Glu	Glu	Lys	Leu	Arg	Lys	Thr	Glu	Glu
				425					430					435
Ile	Ala	Gln	Glu	Arg	Gln	Lys	Gln	Leu	Glu	Ser	Leu	Gly	Ile	Ser
				440					445					450
Leu	Gln	Ser	Ser	Gly	Ile	Lys	Val	Gly	Asp	Asp	Lys	Cys	Phe	Leu
				455					460					465
Val	Asn	Leu	Asn	Ala	Asp	Pro	Ala	Leu	Asn	Glu	Leu	Leu	Val	Tyr
				470					475					480
Tyr	Leu	Lys	Glu	His	Thr	Leu	Ile	Gly	Ser	Ala	Asn	Ser	Gln	Asp
				485					490					495
Ile	Gln	Leu	Cys	Gly	Met	Gly	Ile	Leu	Pro	Glu	His	Cys	Ile	Ile
				500					505					510
Asp	Ile	Thr	Ser	Glu	Gly	Gln	Val	Met	Leu	Thr	Pro	Gln	Lys	Asn
				515					520					525
Thr	Arg	Thr	Phe	Val	Asn	Gly	Ser	Ser	Val	Ser	Ser	Pro	Ile	Gln
				530					535					540
Leu	His	His	Gly	Asp	Arg	Ile	Leu	Trp	Gly	Asn	Asn	His	Phe	Phe
				545					550					555
Arg	Leu	Asn	Leu	Pro	Lys	Lys	Lys	Lys	Lys	Ala	Glu	Arg	Glu	Asp
				560					565					570
Glu	Asp	Gln	Asp	Pro	Ser	Met	Lys	Asn	Glu	Asn	Ser	Ser	Glu	Gln
				575					580					585
Leu	Asp	Val	Asp	Gly	Asp	Ser	Ser	Ser	Glu	Val	Ser	Ser	Glu	Val
				590					595					600
Asn	Phe	Asn	Tyr	Glu	Tyr	Ala	Gln	Met	Glu	Val	Thr	Met	Lys	Ala
				605					610					615
Leu	Gly	Ser	Asn	Asp	Pro	Met	Gln	Ser	Ile	Leu	Asn	Ser	Leu	Glu
				620					625					630
Gln	Gln	His	Glu	Glu	Glu	Lys	Arg	Ser	Ala	Leu	Glu	Arg	Gln	Arg
				635					640					645
Leu	Met	Tyr	Glu	His	Glu	Leu	Glu	Gln	Leu	Arg	Arg	Arg	Leu	Ser
				650					655					660
Pro	Glu	Lys	Gln	Asn	Cys	Arg	Ser	Met	Asp	Arg	Phe	Ser	Phe	His
				665					670					675
Ser	Pro	Ser	Ala	Gln	Gln	Arg	Leu	Arg	Gln	Trp	Ala	Glu	Glu	Arg
				680					685					690
Glu	Ala	Thr	Leu	Asn	Asn	Ser	Leu	Met	Arg	Leu	Arg	Glu	Gln	Ile
				695					700					705
Val	Lys	Ala	Asn	Leu	Leu	Val	Arg	Glu	Ala	Asn	Tyr	Ile	Ala	Glu
				710					715					720
Glu	Leu	Asp	Lys	Arg	Thr	Glu	Tyr	Lys	Val	Thr	Leu	Gln	Ile	Pro

	725		730		735
Ala Ser Ser Leu	Asp Ala Asn Arg Lys	Arg Gly Ser Leu Leu	Ser		
	740		745		750
Glu Pro Ala Ile	Gln Val Lys Thr Lys	Lys Lys Arg			
	755		760		

<210> 458  
 <211> 255  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:450798.1.orf3:2000SEP08

<400> 458

Pro Pro Ser Ser	Pro Arg Ser Pro Tyr Ser	Gly Arg Pro Thr Arg	
1	5	10	15
Phe Cys Glu Leu	Pro Ser His Leu Gln Leu	Pro Ala Leu Asp Pro	
	20	25	30
Ala Met Arg Glu	Ile Leu His Ile Gln Gly	Gly Gln Cys Gly Asn	
	35	40	45
Gln Ile Gly Ser	Lys Phe Trp Glu Val Val	Cys Asp Glu His Gly	
	50	55	60
Ile Asp Pro Thr	Gly Arg Tyr Val Gly Thr	Ser Asp Leu Gln Leu	
	65	70	75
Glu Arg Val Asn	Val Tyr Tyr Asn Glu Ala	Ser Cys Gly Pro Ala	
	80	85	90
Ser Cys Arg Ala	Leu Cys Ser Trp Thr Ser	Ser Leu Ala Pro Trp	
	95	100	105
Thr Ala Ser Ala	Pro Asp Arg Thr Ala Arg	Ser Ser Ala Pro Thr	
	110	115	120
Thr Leu Ser Ser	Asp Ser Leu Ala Ala Gly	Asn Asn Trp Ala Gln	
	125	130	135
Gly Pro Leu His	Arg Gly Arg Arg Ala His	Leu Thr Pro Cys Leu	
	140	145	150
Thr Leu Ser Glu	Arg Arg Leu Arg Thr Ala	Thr Ala Phe Gln Gly	
	155	160	165
Phe Gln Val Cys	His Val Pro Trp Arg Arg	His Trp Ile Trp Asn	
	170	175	180
Gly His Ala Phe	Ser Leu Ser Lys Ile Arg	Glu Asp Val Pro Leu	
	185	190	195
Ile Gly Met Thr	Ala Asp Ser Cys Val Cys	Leu Val Arg Ser Pro	
	200	205	210
Lys Val Ser Asp	Thr Cys Cys Leu Ser His	Thr Asn Ala Thr Pro	
	215	220	225
Val Cys Pro Ser	Gly Trp Leu Ser Tyr Ala	Asp Val Met His Gly	
	230	235	240
Ser Trp Thr Tyr	Asp Ser Thr Phe Met Thr	Phe Ala Ser Glu Leu	
	245	250	255

<210> 459  
 <211> 156  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:410317.1.orf1:2000SEP08

<400> 459

Pro Gln Glu Ile	Asp Phe Leu Gln Gln	Leu Tyr Glu Met Glu	Leu
-----------------	---------------------	---------------------	-----

1	5	10	15
Ser Gln Val Gln Thr His Val Ser Asn Thr Asn Val Val Leu Ser			
20	25	30	
Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Ala Glu			
35	40	45	
Val Lys Ala Gln Tyr Glu Leu Ile Ala Gln Arg Ser Arg Ala Glu			
50	55	60	
Ala Glu Ala Trp Tyr Gln Thr Lys Tyr Glu Glu Leu Gln Val Thr			
65	70	75	
Ala Gly Lys His Gly Asp Asn His Ala Gly His Gln Glu Arg Asp			
80	85	90	
Leu Leu Ser Ser Pro Ala Thr Ile Gln Arg Leu Gln Gly Glu Ala			
95	100	105	
Asp Ala Ala Lys Lys Gln Cys Gln Gln Leu Gln Thr Ala Ile Ala			
110	115	120	
Glu Arg Gly Ala Ala Trp Gly Ala Gly Thr Gln Gly Cys Ser Glu			
125	130	135	
Glu Ala Arg Gly Ser Gly Cys Gly Pro Ala Pro Gly Gln Gly Gly			
140	145	150	
Pro Asp Thr Ala Ala Ala			
155			

&lt;210&gt; 460

&lt;211&gt; 216

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:340268.1.orf3:2000SEP08

&lt;400&gt; 460

Leu Gly Leu Arg Pro Ser Thr Pro Gln Gln Gly Thr Gln Pro Ser		
1	5	10
Arg Leu Cys Cys Pro Cys Thr Pro Leu Val Thr Pro Leu Val Leu		
20	25	30
Ser Tyr Thr Ala Glu Thr His Thr Val Pro Val Tyr Lys Gly Tyr		
35	40	45
Ala Phe Pro His Ala Ile Leu Cys Leu Asp Asp Trp His Leu Thr		
50	55	60
Asn His Pro Met Glu Ile Phe Arg Glu Pro Ser Tyr Cys Phe Thr		
65	70	75
Thr Thr Ala Lys Arg Glu Ile Leu Cys Asn Ile Lys Glu Lys Leu		
80	85	90
Arg Lys Trp Pro Pro Gln Arg Pro Pro Pro Pro Trp Arg Arg Ser		
95	100	105
Arg Ser Cys Pro Gln Trp Phe Gln Cys Leu Glu Val Leu Phe Gln		
110	115	120
Ala Ser Phe Leu Gly Met Glu Ser Cys Trp His Pro Arg Asp His		
125	130	135
Leu Gln Leu His His Glu Cys Cys Asp Cys Gly His Pro Ala Lys		
140	145	150
Thr Cys Thr Pro Thr Arg Met Leu Ser Trp Arg His His His Val		
155	160	165
Pro Trp Ala Leu Pro Thr Gly Cys Arg Arg Ser Arg Pro Trp Cys		
170	175	180
Pro Ala Pro Arg Arg Ser Ser Pro Ser Leu Ser Ala Ser Pro Cys		
185	190	195
Val Glu Leu Ala Ala Pro Leu Leu Ala Ala Phe Gln Gln Asn Val		
200	205	210
Glu Ser Ser Ser Thr Glu		
215		

<210> 461  
 <211> 160  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:2051671.1.orf3:2000SEP08

<400> 461  
 Thr Gln Ser Ile Met Asp Val Gly Trp Pro Glu Leu His Ala Pro  
 1 5 10 15  
 Pro Leu Asp Lys Met Cys Thr Ile Cys Lys Ala Gln Glu Ser Trp  
 20 25 30  
 Leu Asn Ser Asn Leu Gln His Val Val Val Ile His Cys Arg Gly  
 35 40 45  
 Gly Lys Gly Arg Ile Gly Val Val Ile Ser Ser Tyr Met His Phe  
 50 55 60  
 Thr Asn Val Ser Ala Ser Ala Asp Gln Ala Leu Asp Arg Phe Ala  
 65 70 75  
 Met Lys Lys Phe Tyr Asp Asp Lys Val Ser Ala Leu Met Gln Pro  
 80 85 90  
 Ser Gln Lys Thr Val Cys Ser Val Pro Gln Trp Ala Pro Val Arg  
 95 100 105  
 Ile Gly Glu Asn Glu Cys Leu Ser Pro Val Pro Ala Phe Cys His  
 110 115 120  
 Pro Pro Gly Pro Leu Cys Arg Ala Gly Ala Glu Gln Arg Ser Cys  
 125 130 135  
 Phe Val Pro Glu Val Gln Arg Asp Gln Pro Gln Asn Thr Pro Ser  
 140 145 150  
 Ser Pro Gly Arg Thr Gln Leu Ser Gly Gly  
 155 160

<210> 462  
 <211> 97  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID, No: LG:998844.1.orf3:2000SEP08

<400> 462  
 Ala Glu Glu Pro Gly Pro Gly Arg Ala Gly Thr Thr Pro Glu Ser  
 1 5 10 15  
 Gly Gly Ser Ser Gly Gly Gly Arg Gly Ser Leu Ser Ser Gln Arg  
 20 25 30  
 Arg Ser Asp Ala Ala Gly Thr Met Gly Cys Cys Thr Gly Arg Cys  
 35 40 45  
 Ser Leu Ile Cys Leu Cys Ala Leu Gln Leu Val Ser Ala Leu Glu  
 50 55 60  
 Arg Gln Ile Phe Asp Phe Leu Gly Phe Gln Trp Ala Pro Ile Leu  
 65 70 75  
 Gly Asn Phe Leu His Ile Ile Val Val Ile Leu Gly Leu Phe Gly  
 80 85 90  
 Thr Ile Gln His Ile Pro Leu  
 95

<210> 463  
 <211> 124  
 <212> PRT  
 <213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1043787.1.orf1:2000SEP08

&lt;400&gt; 463

Gly	Lys	Thr	Asp	Val	Asn	Tyr	Thr	Gln	Leu	Val	Asp	Leu	His	Ala	
1				5					10					15	
Arg	Tyr	Ala	Glu	Cys	Gly	Leu	Arg	Ile	Leu	Ala	Phe	Pro	Cys	Asn	
				20					25					30	
Gln	Phe	Gly	Arg	Gln	Glu	Pro	Gly	Ser	Asn	Gln	Glu	Ile	Lys	Glu	
				35					40					45	
Phe	Ala	Ala	Gly	Tyr	Asn	Val	Arg	Phe	Asp	Met	Tyr	Ser	Lys	Ile	
				50					55					60	
Cys	Val	Asn	Gly	Asp	Asp	Ala	His	Pro	Leu	Trp	Lys	Trp	Met	Lys	
				65					70					75	
Val	Gln	Pro	Lys	Gly	Arg	Gly	Met	Leu	Gly	Asn	Ala	Ile	Lys	Trp	
				80					85					90	
Asn	Phe	Thr	Lys	Phe	Leu	Ile	Asp	Lys	Asn	Gly	Cys	Val	Val	Lys	
				95					100					105	
Arg	Tyr	Gly	Pro	Met	Glu	Glu	Pro	Gln	Val	Ile	Glu	Lys	Asp	Leu	
				110					115					120	
Pro	Cys	Tyr	Leu												

&lt;210&gt; 464

&lt;211&gt; 68

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1098931.16.orf1:2000SEP08

&lt;400&gt; 464

His	Val	Ile	Asn	Ile	His	Pro	Tyr	Asp	Val	His	Ile	Lys	Lys	Lys	
1				5					10					15	
Lys	Thr	Gly	Ala	Phe	Ser	Ala	Gly	Asn	Trp	Ile	Arg	Ser	Leu	Thr	
				20					25					30	
Lys	Val	Phe	Phe	Lys	Gly	Phe	Lys	Tyr	Leu	Tyr	Leu	Thr	Pro	Gln	
				35					40					45	
Asp	Tyr	Thr	Arg	Ile	Ser	Ser	Leu	Asn	Ser	Val	His	Cys	Lys	His	
				50					55					60	
Ile	Glu	Glu	Gly	Gly	Glu	Ser	Arg								
				65											

&lt;210&gt; 465

&lt;211&gt; 85

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:199423.2.orf1:2000SEP08

&lt;400&gt; 465

Leu	Arg	Ala	Gly	Arg	Ile	Val	Ser	Cys	Leu	Gly	Leu	Gly	Cys	Ser	
1				5					10					15	
Trp	Gln	Tyr	Phe	Ile	Leu	Leu	Ile	Ile	Thr	Asp	Gly	Val	Ile	Ser	
				20					25					30	
Asp	Met	Glu	Glu	Thr	Arg	His	Ala	Val	Val	Gln	Ala	Ser	Lys	Leu	
				35					40					45	
Pro	Met	Ser	Ile	Ile	Ile	Val	Gly	Val	Gly	Asn	Ala	Asp	Phe	Ala	
				50					55					60	



Ala Met Glu Phe Leu Asp Gly Asp Ser Arg Met Leu Arg Ser His  
                                 65                                75  
 Thr Gly Glu Glu Ala Ala Arg Asp Ile Val  
                                 80                                85

<210> 466

<211> 172

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1075297.1.orf2:2000SEP08

<400> 466

Arg Phe Leu Leu Gly Phe Ser His Leu Lys Ile Glu Thr Lys Ile  
   1                                5                                10                                15  
 Glu Ser Met Ala Asp Leu Lys Gln Leu Met Asp Asn Glu Val Leu  
                                 20                                25                                30  
 Met Ala Phe Thr Ser Tyr Ala Thr Ile Ile Leu Ala Lys Met Met  
                                 35                                40                                45  
 Phe Leu Ser Ser Ala Thr Ala Phe Gln Arg Leu Thr Asn Lys Val  
                                 50                                55                                60  
 Phe Ala Asn Pro Glu Asp Cys Ala Gly Phe Gly Lys Gly Glu Asn  
                                 65                                70                                75  
 Ala Lys Lys Phe Leu Arg Thr Asp Glu Lys Val Glu Arg Val Arg  
                                 80                                85                                90  
 Arg Ala His Leu Asn Asp Leu Glu Asn Ile Val Pro Phe Leu Gly  
                                 95                                100                                105  
 Ile Gly Leu Leu Tyr Ser Leu Ser Gly Pro Asp Leu Ser Thr Ala  
                                 110                                115                                120  
 Leu Ile His Phe Arg Ile Phe Val Gly Ala Arg Ile Tyr His Thr  
                                 125                                130                                135  
 Ile Ala Tyr Leu Thr Pro Leu Pro Gln Pro Asn Arg Gly Leu Ala  
                                 140                                145                                150  
 Phe Phe Val Gly Tyr Gly Val Thr Leu Ser Met Ala Tyr Arg Leu  
                                 155                                160                                165  
 Leu Arg Ser Arg Leu Tyr Leu  
                                 170

<210> 467

<211> 114

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1043321.1.orf3:2000SEP08

<400> 467

Glu Leu Val Val Gly Ala Val Ile Pro Gly Pro Phe Val His Ile  
   1                                5                                10                                15  
 Val Leu Glu Cys His Gly Val Gly Asp Asp Glu Lys Glu Leu Thr  
                                 20                                25                                30  
 Gly Arg Ser Ser Ser Leu Val Ser Pro Val Ala Pro Gln Ala Met  
                                 35                                40                                45  
 Ser Pro Gly Cys Asp Ala Gln Gly Ser Asp Tyr Val Val Gln Ile  
                                 50                                55                                60  
 Cys Lys Glu Ala Gly Val Asp Leu Ala Gly Gly Asp Gln Cys Asp  
                                 65                                70                                75  
 Pro Val Gln Gly Ser Asn Val Gln Gln Gly His Glu Asp Asp Val  
                                 80                                85                                90  
 Leu Ser Arg Arg Thr Pro Ala Trp Ala Gly Ala Pro Pro His Pro

95 100 105  
 Gly Ser Trp Gly Arg Ile Gly His Leu  
 110

<210> 468  
 <211> 110  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:297070.1.orf1:2000SEP08

<400> 468  
 Met Met Gly Ser Trp Lys His Cys Leu Phe Ser Ala Ser Leu Ile  
 1 5 10 15  
 Ser Ala Leu Ile Phe Val Phe Val Tyr Asn Thr Glu Leu Trp Glu  
 20 25 30  
 Asn Lys Arg Phe Leu Arg Ala Ala Leu Ser Asn Ala Ser Leu Leu  
 35 40 45  
 Ala Glu Ala Cys His Gln Ile Phe Glu Gly Lys Val Phe Tyr Pro  
 50 55 60  
 Thr Glu Asn Ala Leu Lys Thr Thr Leu Asp Glu Ala Thr Cys Tyr  
 65 70 75  
 Glu Tyr Met Val Arg Ser His Tyr Val Thr Glu Thr Leu Ser Glu  
 80 85 90  
 Glu Glu Ala Gly Phe Pro Leu Ala Tyr Thr Val Thr Ile His Lys  
 95 100 105  
 Asp Phe Arg His Phe  
 110

<210> 469  
 <211> 212  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1085041.1.orf3:2000SEP08

<400> 469  
 Cys Phe Ala His Val Ala Gly Leu Arg Leu Arg His Leu Leu Tyr  
 1 5 10 15  
 Ser Lys Gln Met Ala Ala Leu Pro Met Leu Trp Thr Gly Leu Val  
 20 25 30  
 Leu Leu Gly Leu Leu Gly Phe Pro Gln Thr Pro Ala Gln Gly His  
 35 40 45  
 Asp Thr Val Gln Pro Asn Phe Gln Gln Asp Lys Phe Leu Gly Arg  
 50 55 60  
 Trp Tyr Ser Ala Gly Leu Ala Ser Asn Ser Ser Trp Phe Arg Glu  
 65 70 75  
 Lys Lys Glu Leu Leu Phe Met Cys Gln Thr Val Val Ala Pro Ser  
 80 85 90  
 Thr Glu Gly Gly Leu Asn Leu Thr Ser Thr Phe Leu Arg Lys Asn  
 95 100 105  
 Gln Cys Glu Thr Lys Val Met Val Leu Gln Pro Ala Gly Val Pro  
 110 115 120  
 Gly Gln Tyr Thr Tyr Asn Ser Pro His Trp Gly Ser Phe His Ser  
 125 130 135  
 Leu Ser Val Val Glu Thr Asp Tyr Asp Glu Tyr Ala Phe Leu Phe  
 140 145 150  
 Ser Lys Gly Thr Lys Gly Pro Gly Gln Asp Phe Arg Met Ala Thr  
 155 160 165

Leu Tyr Ser Arg Ala Gln Leu Leu Lys Glu Glu Leu Lys Glu Lys  
 170 175 180  
 Phe Ile Thr Phe Ser Lys Asp Gln Gly Leu Thr Glu Glu Asp Ile  
 185 190 195  
 Val Phe Leu Pro Gln Pro Asp Lys Cys Ile Thr Arg Val Asn Thr  
 200 205 210  
 Gly Glu

<210> 470  
 <211> 178  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1071544.1.orf2:2000SEP08

<400> 470  
 Trp Arg Leu Ala Val Pro Pro Arg Ala Pro Pro Leu Arg Ala Pro  
 1 5 10 15  
 Ala Gly Gly Arg Ser Arg Leu Trp Phe Ser Ser Leu Ser Ala Gly  
 20 25 30  
 Thr Ser Arg Pro Thr Ser Phe Arg Thr Met Ala Asn Leu Glu Arg  
 35 40 45  
 Thr Phe Ile Ala Ile Lys Pro Asp Gly Val Gln Arg Gly Leu Val  
 50 55 60  
 Gly Glu Ile Ile Lys Arg Phe Glu Gln Lys Gly Phe Arg Leu Val  
 65 70 75  
 Ala Met Lys Phe Leu Arg Ala Ser Glu Glu His Leu Lys Gln His  
 80 85 90  
 Tyr Ile Asp Leu Lys Asp Arg Pro Phe Phe Pro Gly Leu Val Lys  
 95 100 105  
 Tyr Met Asn Ser Gly Pro Val Val Ala Met Val Trp Glu Gly Leu  
 110 115 120  
 Asn Val Val Lys Thr Gly Arg Val Met Leu Gly Glu Thr Asn Pro  
 125 130 135  
 Ala Asp Ser Lys Pro Gly Thr Ile Arg Gly Gly Phe Leu Ala Phe  
 140 145 150  
 Lys Leu Ala Gly Thr Ser Phe Thr Ala Val Ile Gln Trp Arg Val  
 155 160 165  
 Pro Arg Lys Arg Ser Gly Tyr Gly Leu Ser Pro Lys Asn  
 170 175

<210> 471  
 <211> 152  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:2052480.1.orf2:2000SEP08

<400> 471  
 Gly Gln Asp Arg Asn Thr Ala Leu Trp Leu Pro Trp Ser Lys Tyr  
 1 5 10 15  
 Lys Lys Lys Ser Leu Ala Asn Thr Arg Glu Ile Arg Gly Gly Pro  
 20 25 30  
 Trp Val Trp Asn Arg Ser Glu Ala Asp Cys Ile Ala Val Gln His  
 35 40 45  
 Val Cys Thr Ile Val Ser Phe Arg Ser Ala Asn Leu Cys Ala Ala  
 50 55 60  
 Ala Leu Ala Ala Ile Leu Thr Arg Leu Arg Glu Asn Asn Glu Gly

	65		70		75
Gly Thr Ala Pro Asp His Ser Gly His Gly Arg His Pro Leu Gln					
	80		85		90
Asp Ile Thr Leu Ser Thr Gln Asn Ala Cys Thr Arg Val Val Arg					
	95		100		105
Lys Leu Val Pro Ser Cys Asp Val Arg Phe Leu Leu Ser Glu Ser					
	110		115		120
Gly Ser Thr Lys Gly Ala Ala Met Val Thr Ala Val Ala Ser Pro					
	125		130		135
Arg Ala Gly Pro Ala Glu Ala Asp Arg Gln Gly Ala Gly Phe Val					
	140		145		150
Pro Ser					

&lt;210&gt; 472

&lt;211&gt; 164

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:450105.1.orf2:2000SEP08

&lt;400&gt; 472

Pro Ser Ala Ala Arg Pro Pro Tyr Ser Arg Tyr Arg Ala Arg Arg			
1 5 10 15			
His His Leu Arg Arg Ser Asn Met Gly Lys Thr Arg Gly Met Gly			
20 25 30			
Ala Gly Arg Lys Leu Lys Thr His Arg Arg Asn Gln Arg Trp Ala			
35 40 45			
Asp Lys Ala Tyr Lys Lys Ser His Leu Gly Asn Glu Trp Lys Lys			
50 55 60			
Pro Phe Ala Gly Ser Ser His Ala Lys Gly Ile Val Leu Glu Lys			
65 70 75			
Ile Gly Ile Glu Ala Lys Gln Pro Asn Ser Ala Ile Arg Lys Cys			
80 85 90			
Ala Arg Val Gln Leu Val Lys Asn Gly Lys Lys Ile Ala Ala Phe			
95 100 105			
Val Pro Asn Asp Gly Cys Leu Asn Tyr Ile Glu Glu Asn Asp Glu			
110 115 120			
Val Leu Ile Ala Gly Phe Gly Arg Lys Gly His Ala Val Gly Asp			
125 130 135			
Ile Pro Gly Val Arg Phe Lys Val Val Lys Val Ser Gly Val Ser			
140 145 150			
Leu Leu Ala Leu Phe Lys Glu Lys Lys Glu Lys Pro Arg Ser			
155 160			

&lt;210&gt; 473

&lt;211&gt; 100

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:450581.1.orf2:2000SEP08

&lt;400&gt; 473

Gly Met Ser Arg Gly Pro Ser Ala Arg Val Thr Ser Ser Pro Cys			
1 5 10 15			
Ser Ser Pro Arg Gly Arg Pro Gly Gly Cys Ala Glu Ala Leu Ala			
20 25 30			
Phe Leu Val Gly Ser Cys Leu Val Phe Gly Val Phe Glu Phe Val			
35 40 45			

Leu	Cys	Leu	Asn	Phe	Cys	Asp	Pro	Val	Leu	Arg	Val	Val	Ser	Leu	
				50					55					60	
Val	Ile	Arg	Leu	Gln	Phe	Cys	Leu	Ser	Trp	Leu	Lys	Met	Glu	Leu	
				65					70					75	
Val	Leu	Arg	His	Gly	Tyr	Pro	Cys	Tyr	Ile	Ile	Tyr	Lys	Thr	Thr	
				80					85					90	
His	Ala	Leu	Asn	Thr	Lys	Lys	Lys	Lys	Lys						
				95					100						

&lt;210&gt; 474

&lt;211&gt; 122

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:450887.1.orf3:2000SEP08

&lt;400&gt; 474

Ser	Val	His	Phe	Ser	Arg	Lys	Gly	Phe	Val	Leu	Met	Ala	Pro	Pro	
1				5					10					15	
Gln	Pro	Lys	Ser	Gly	Leu	Phe	Val	Gly	Ile	Asn	Lys	Gly	His	Val	
				20					25					30	
Val	Thr	Lys	Arg	Glu	Leu	Pro	Pro	Arg	Pro	Cys	His	Arg	Lys	Gly	
				35					40					45	
Lys	Ser	Thr	Lys	Arg	Val	Ser	Met	Val	Arg	Gly	Leu	Ile	Arg	Glu	
				50					55					60	
Val	Ala	Gly	Phe	Ala	Pro	Tyr	Glu	Lys	Arg	Ile	Thr	Glu	Leu	Leu	
				65					70					75	
Lys	Val	Gly	Lys	Asp	Lys	Arg	Ala	Leu	Lys	Leu	Ala	Lys	Arg	Lys	
				80					85					90	
Leu	Gly	Thr	His	Lys	Arg	Ala	Lys	Lys	Lys	Arg	Glu	Glu	Met	Ala	
				95					100					105	
Gly	Val	Leu	Arg	Lys	Met	Arg	Ser	Ala	Gly	Thr	His	Thr	Asp	Lys	
				110					115					120	
Lys	Lys														

&lt;210&gt; 475

&lt;211&gt; 101

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:460809.1.orf3:2000SEP08

&lt;400&gt; 475

Ala	Trp	Val	Glu	Trp	Ala	Ser	Arg	Ser	Ala	Pro	Arg	Ala	His	Arg	
1				5					10					15	
Glu	Ile	Gln	Lys	Phe	Ala	Met	Lys	Glu	Met	Gly	Thr	Pro	Asn	Leu	
				20					25					30	
His	Ile	Asp	Val	Arg	Leu	Asn	Lys	Ala	Leu	Trp	Ala	Lys	Gly	Ile	
				35					40					45	
Arg	Asn	Val	Pro	Tyr	His	Ile	His	Met	Lys	Leu	Pro	Arg	Lys	Leu	
				50					55					60	
Asn	Glu	Asp	Glu	Asp	Ser	Pro	Asp	Lys	Leu	Tyr	Ala	Leu	Val	Pro	
				65					70					75	
Thr	Tyr	Thr	Cys	Tyr	His	Phe	His	Lys	Ser	Ile	Asp	Arg	Gln	Cys	
				80					85					90	
Gly	Arg	Glu	Leu	Thr	Thr	Asp	Gly	Ser	Ile	His					
				95					100						

<210> 476  
 <211> 207  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:452089.1.orf2:2000SEP08

<400> 476  
 Leu Gly Lys His Arg Arg Pro Pro Pro Pro Lys Asp Gly Arg Arg  
   1                  5                  10                  15  
 Arg Gly Tyr Gly Arg Leu Leu Ala Pro Met Ala His Glu Lys Lys  
                   20                  25                  30  
 Leu Ser Asn Pro Met Arg Glu Ile Lys Val Gln Lys Leu Val Leu  
                   35                  40                  45  
 Asn Ile Ser Val Gly Glu Ser Gly Asp Arg Leu Thr Arg Ala Ala  
                   50                  55                  60  
 Lys Val Leu Glu Gln Leu Ser Gly Gln Thr Pro Val Phe Ser Lys  
                   65                  70                  75  
 Ala Arg Tyr Thr Val Arg Ser Phe Gly Ile Arg Arg Asn Glu Lys  
                   80                  85                  90  
 Ile Ala Cys Tyr Val Thr Val Arg Gly Glu Lys Ala Met Gln Leu  
                   95                  100                 105  
 Leu Glu Ser Gly Leu Lys Val Lys Glu Tyr Glu Leu Leu Arg Arg  
                  110                 115                 120  
 Asn Phe Ser Asp Thr Gly Cys Phe Gly Phe Gly Ile Gln Glu His  
                  125                 130                 135  
 Ile Asp Leu Gly Ile Lys Tyr Asp Pro Ser Thr Gly Ile Tyr Gly  
                  140                 145                 150  
 Met Asp Phe Tyr Val Val Leu Glu Arg Ala Gly Tyr Arg Val Ala  
                  155                 160                 165  
 Arg Arg Arg Arg Cys Lys Ser Arg Val Gly Ile Gln His Arg Val  
                  170                 175                 180  
 Thr Lys Glu Asp Ser Met Lys Trp Leu Pro Gly Thr Ser Thr Lys  
                  185                 190                 195  
 Ala Ser Ser Ser Asn Lys Ala Gln Ala Asn Thr Leu  
                  200                 205

<210> 477  
 <211> 83  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1099416.1.orf3:2000SEP08

<400> 477  
 Gly Cys Leu Ala Gly Ile Arg Lys Asp Asn Lys Met Lys Gly Thr  
   1                  5                  10                  15  
 Ser Pro Phe Gly Lys Cys Arg Asp Met Ile His Lys Leu Cys Cys  
                   20                  25                  30  
 Leu Cys Gly Ser Lys Ala Tyr His Leu Gln Lys Ser Thr Cys Gly  
                   35                  40                  45  
 Lys Cys Gly Ser Pro Ala Lys Arg Lys Arg Lys Cys Asn Trp Thr  
                   50                  55                  60  
 Ala Thr Ala Lys Arg Lys Tyr His Gly Asp Trp Leu Asn Glu Ala  
                   65                  70                  75  
 Pro Lys His Cys Ile Leu Gln Ile  
                   80

<210> 478

<211> 75  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:255713.1.orf1:2000SEP08

<400> 478  
 Lys Thr Pro Asp Met Cys Ile Asp Thr Arg Leu Asn Lys Ala Val  
           1                  5                  10                  15  
 Trp Ala Lys Val Val Arg Asn Val His Thr Val Ser Met Cys Ser  
                   20                  25                  30  
 Gly Trp Ser Arg Lys His Asn Glu Val Arg Asn Ser Pro Asn Lys  
                   35                  40                  45  
 Leu Tyr Thr Leu Val Ile Tyr Leu Leu Pro Leu Ser Gln Ile Tyr  
                   50                  55                  60  
 Ser Gln Cys Gly Trp Glu Leu Thr Ala Asp His Gln Ile His Gln  
                   65                  70                  75

<210> 479  
 <211> 162  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:998903.1.orf1:2000SEP08

<400> 479  
 Ala Arg Cys Ser Leu Ser Ala Asp Leu Pro Ser Gln Pro Pro Arg  
           1                  5                  10                  15  
 Asn Pro Phe Thr Ala Arg Lys Met Ala Glu Gln Glu Ala Pro Val  
                   20                  25                  30  
 Ala Val Glu Ala Pro Thr Pro Val Leu Gly Glu Pro Met Asp Leu  
                   35                  40                  45  
 Met Thr Ala Leu Gln Leu Val Met Lys Lys Ser Ser Ala His Asp  
                   50                  55                  60  
 Gly Leu Val Lys Gly Leu Arg Glu Ala Ala Lys Ala Ile Glu Lys  
                   65                  70                  75  
 His Ala Ala Gln Leu Cys Val Leu Ala Glu Asp Cys Asp Gln Pro  
                   80                  85                  90  
 Asp Tyr Val Lys Leu Val Lys Ala Leu Cys Ala Glu His Asn Val  
                   95                  100                  105  
 His Leu Val Thr Val Pro Ala Ala Lys Thr Leu Gly Glu Trp Ala  
                   110                  115                  120  
 Gly Leu Cys Lys Ile Asp Ser Glu Gly Lys Ala Arg Lys Val Val  
                   125                  130                  135  
 Gly Cys Ser Cys Val Val Val Lys Asp Tyr Gly Glu Glu Ser Glu  
                   140                  145                  150  
 Gly Leu Asn Ile Val Gln Glu Tyr Val Lys Ser His  
                   155                  160

<210> 480  
 <211> 90  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1119656.1.orf2:2000SEP08

&lt;400&gt; 480

```

Asn Phe Asn Val Arg Lys Leu Lys Glu Lys Met Asp Phe Val Leu
 1           5           10           15
Cys Leu Pro Leu Leu Val Leu Ser Val Tyr Ser Tyr Ile Asp Ser
           20           25           30
Val Val Ala Met His Asn Phe Pro Thr Asn Phe Ile Leu Thr Cys
           35           40           45
Arg Glu Ser Val Asp Lys Ile Phe Cys Asn Lys Val Leu Phe Ala
           50           55           60
Asn Met Tyr Phe Ile Phe Thr Val Tyr Ser Ile Phe Leu Ile Pro
           65           70           75
Tyr Lys Phe Leu Gln Glu Ser Phe Arg Phe Ser Ile Gln Asn Gly
           80           85           90

```

&lt;210&gt; 481

&lt;211&gt; 108

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1096907.1.orf1:2000SEP08

&lt;400&gt; 481

```

Trp Ser Phe Gly Ser Gly Asp Met Ala Lys Arg Thr Lys Lys Val
 1           5           10           15
Gly Ile Val Gly Lys Tyr Gly Thr Arg Tyr Gly Ala Ser Leu Arg
           20           25           30
Lys Met Val Lys Lys Ile Glu Ile Ser Gln His Ala Lys Tyr Thr
           35           40           45
Cys Ser Phe Cys Gly Lys Thr Lys Met Lys Arg Arg Ala Val Gly
           50           55           60
Ile Trp His Cys Gly Ser Cys Met Lys Thr Val Ala Gly Gly Ala
           65           70           75
Trp Thr Tyr Asn Thr Thr Ser Ala Val Thr Val Lys Ser Ala Ile
           80           85           90
Arg Arg Leu Lys Ala Gly Glu Phe Val Trp Trp Trp Arg Glu Arg
           95          100          105
Gly Glu Val

```

&lt;210&gt; 482

&lt;211&gt; 142

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1323741.1.orf2:2000SEP08

&lt;400&gt; 482

```

Ile Tyr Ala Leu Ser Leu Gly Ala Gly Gly Ala Ala Ala Ser Ala
 1           5           10           15
Ala Cys Ala Ala Met Ala Lys Ile Lys Ala Arg Asp Leu Arg Gly
           20           25           30
Lys Lys Lys Glu Glu Leu Leu Lys Gln Leu Asp Asp Leu Lys Val
           35           40           45
Glu Leu Ser Gln Leu Arg Val Ala Lys Val Thr Gly Gly Ala Ala
           50           55           60
Ser Lys Leu Ser Lys Ile Arg Val Val Arg Lys Ser Ile Ala Arg
           65           70           75
Val Leu Thr Val Ile Asn Gln Thr Gln Lys Glu Asn Leu Arg Lys

```



	80		85		90
Phe Tyr Lys Gly Lys Lys Tyr Lys Pro Leu Asp Leu Arg Pro Lys					
	95		100		105
Lys Thr Arg Ala Met Arg Arg Arg Leu Thr Lys His Glu Glu Lys					
	110		115		120
Leu Lys Thr Lys Lys Gln Gln Arg Lys Glu Arg Leu Tyr Pro Leu					
	125		130		135
Arg Lys Tyr Ala Val Lys Ala					
	140				

&lt;210&gt; 483

&lt;211&gt; 82

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1098372.1.orf1:2000SEP08

&lt;400&gt; 483

Lys Lys Arg Lys Gln Val Pro Lys Phe Thr Leu Asp Arg Thr His					
1	5		10		15
Pro Val Glu Asp Gly Ile Met Asp Ala Ala Asn Phe Glu Gln Phe					
	20		25		30
Phe Gln Glu Arg Ile Lys Met Asn Gly Lys Ala Gly Asn Phe Gly					
	35		40		45
Gly Gly Val Val Thr Ile Glu Gly Ser Lys Ser Lys Thr Ser Val					
	50		55		60
Thr Ser Lys Leu Pro Phe Ser Asn Arg Tyr Leu Lys Tyr Leu Thr					
	65		70		75
Lys Lys Ile Ser Glu Glu Glu					
	80				

&lt;210&gt; 484

&lt;211&gt; 163

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:1006783.1.orf1:2000SEP08

&lt;400&gt; 484

Leu Ala Lys Met Thr Asn Thr Lys Gly Lys Arg Arg Gly Thr Arg					
1	5		10		15
Tyr Met Phe Ser Arg Pro Phe Arg Lys His Gly Val Val Pro Leu					
	20		25		30
Ala Thr Tyr Met Arg Ile Tyr Lys Lys Gly Asp Ile Val Asp Ile					
	35		40		45
Lys Gly Met Gly Thr Val Gln Lys Gly Met Pro His Lys Cys Tyr					
	50		55		60
His Gly Lys Thr Gly Arg Val Tyr Asn Val Thr Gln His Ala Val					
	65		70		75
Gly Ile Ile Val Asn Lys Gln Val Lys Gly Lys Ile Leu Ala Lys					
	80		85		90
Arg Ile Asn Val Arg Ile Glu His Ile Lys His Ser Lys Ser Arg					
	95		100		105
Asp Ser Phe Leu Lys Arg Val Lys Glu Asn Asp Gln Lys Lys Lys					
	110		115		120
Glu Ala Lys Glu Lys Gly Thr Trp Val Gln Leu Lys Arg Gln Pro					
	125		130		135
Ala Pro Pro Arg Glu Ala His Phe Val Arg Thr Asn Gly Lys Glu					
	140		145		150

Pro Glu Leu Leu Glu Pro Ile Pro Tyr Glu Phe Met Ala  
 155 160

<210> 485  
 <211> 118  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1097562.1.orf2:2000SEP08

<400> 485  
 Gly Lys Arg Met Val Ala Ala Lys Lys Thr Lys Lys Ser Leu Glu  
 1 5 10 15  
 Ser Ile Asn Ser Arg Leu Gln Leu Val Met Lys Ser Gly Lys Tyr  
 20 25 30  
 Val Leu Gly Tyr Lys Gln Thr Leu Lys Met Ile Arg Gln Gly Lys  
 35 40 45  
 Ala Lys Leu Val Ile Leu Ala Asn Asn Cys Pro Ala Leu Arg Lys  
 50 55 60  
 Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Thr Gly Val His  
 65 70 75  
 His Tyr Ser Gly Asn Asn Ile Glu Leu Gly Thr Ala Cys Gly Lys  
 80 85 90  
 Tyr Tyr Arg Val Cys Thr Leu Ala Ile Ile Asp Pro Gly Asp Ser  
 95 100 105  
 Asp Ile Ile Arg Ser Met Pro Glu Gln Thr Gly Glu Lys  
 110 115

<210> 486  
 <211> 260  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:998868.1.orf2:2000SEP08

<400> 486  
 Met Ala Val Gly Lys Asn Lys Arg Ile Ser Lys Gly Lys Lys Gly  
 1 5 10 15  
 Gly Lys Lys Lys Thr Val Asp Pro Phe Ser Lys Lys Asp Trp Tyr  
 20 25 30  
 Asp Ile Lys Ala Pro Ser Val Phe Ser Val Arg Asn Ile Gly Lys  
 35 40 45  
 Thr Leu Val Ser Arg Thr Gln Gly Thr Arg Ile Ala Ser Glu Gly  
 50 55 60  
 Leu Lys His Arg Val Phe Glu Val Cys Leu Ala Asp Leu Gln Gly  
 65 70 75  
 Asp Glu Asp Gln Ala Tyr Arg Lys Ile Arg Leu Arg Ala Glu Asp  
 80 85 90  
 Val Gln Gly Arg Asn Val Leu Thr Asn Phe Trp Gly Met Asn Phe  
 95 100 105  
 Thr Thr Asp Lys Leu Arg Ser Leu Val Lys Lys Trp Gln Thr Leu  
 110 115 120  
 Ile Glu Ala His Ala Asp Val Lys Thr Thr Asp Asn Tyr Met Leu  
 125 130 135  
 Arg Leu Phe Cys Ile Gly Phe Thr Lys Arg Arg Pro Asn Gln Val  
 140 145 150  
 Lys Arg Thr Cys Tyr Ala Gln Ala Ser Gln Ile Arg Gln Ile Arg  
 155 160 165  
 Arg Lys Met Val Glu Ile Met Ile Asn Gln Ala Ser Thr Cys Asp

170	175	180
Leu Lys Glu Leu Val Ser Lys Phe Ile	Pro Glu Val Ile Gly Lys	
185	190	195
Glu Ile Glu Lys Ser Thr Ser Ser Ile	Phe Pro Leu Gln Asn Val	
200	205	210
Phe Ile Arg Lys Val Lys Ile Leu Lys	Ala Pro Lys Phe Asp Leu	
215	220	225
Gly Lys Leu Met Glu Val His Gly Asp	Tyr Lys Glu Asp Val Gly	
230	235	240
Val Lys Leu Glu Arg Pro Val Glu Gly	Asp Glu Ala Gly Gln Glu	
245	250	255
Val Ala Ala Ala Glu		
260		

<210> 487  
 <211> 115  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1063383.1.orf1:2000SEP08

<400> 487	
Tyr Ser Ala Ser Phe Gly Lys Gly Gly Lys Ala Ser Leu Ala Glu	
1 5 10 15	
Ala Pro Ile Pro Phe Pro Ala Cys Ser Val Ala Gly Asp Trp Lys	
20 25 30	
Thr Leu Cys Asp Ser Val Gln Ser Leu Gly Gly Lys Ala Ser Phe	
35 40 45	
Gln Asp Ser Ser Ser Pro Gly Ala Thr Ser Ser Pro Lys Gly Ile	
50 55 60	
Met Ala Ala Leu Arg Ser Leu Val Lys Pro Lys Ile Val Lys Lys	
65 70 75	
Arg Thr Lys Lys Phe Ile Arg His Gln Ser Asp Arg Tyr Val Lys	
80 85 90	
Ile Lys Arg Asn Trp Trp Lys Pro Arg Gly Ile Asp Asn Arg Val	
95 100 105	
His Arg Arg Phe Lys Gly Gln Val Leu Ile	
110 115	

<210> 488  
 <211> 57  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1400567.1.orf2:2000SEP08

<400> 488	
Met Cys Thr Thr Asp Asn Arg Leu Asn Lys Ala Val Trp Ala Lys	
1 5 10 15	
Gly Ile Arg Asn Val Pro Cys Arg Ile His Val Gln Leu Ser Arg	
20 25 30	
Glu Tyr Asn Glu Asp Glu Asp Ser Pro Asn Lys Leu Cys Thr Ser	
35 40 45	
Val Thr Tyr Val Pro Val Thr Thr Phe Lys Asn Leu	
50 55	

<210> 489  
 <211> 167  
 <212> PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:449404.1.orf1:2000SEP08

&lt;400&gt; 489

```

Ser Gln Pro Lys Ser Cys Leu Arg Ser Gly His Pro Ser Leu His
 1          5          10          15
Ala Thr Met Ser Arg Arg Lys Thr Arg Glu Pro Lys Glu Glu Asn
          20          25          30
Val Thr Leu Gly Pro Thr Val Arg Glu Gly Glu Tyr Val Phe Gly
          35          40          45
Val Ala His Ile Phe Ala Ser Phe Asn Asp Thr Phe Ile His Ile
          50          55          60
Thr Asp Leu Ser Gly Arg Glu Thr Leu Val Arg Ile Thr Gly Gly
          65          70          75
Met Lys Val Lys Ala Asp Arg Asp Glu Ser Ser Pro Tyr Ala Ala
          80          85          90
Met Leu Ala Ala Gln Asp Val Ala Gln Arg Cys Lys Glu Leu Gly
          95          100          105
Ile Thr Ala Leu His Ile Lys Leu Arg Ala Thr Gly Gly Asn Lys
          110          115          120
Thr Lys Thr Pro Gly Pro Gly Ala Gln Ser Ala Leu Arg Ala Leu
          125          130          135
Ala Arg Ser Gly Met Lys Ile Gly Arg Ile Glu Asp Val Thr Pro
          140          145          150
Val Pro Thr Asp Ser Thr Arg Arg Lys Gly Gly Arg Arg Gly Arg
          155          160          165
Arg Leu

```

&lt;210&gt; 490

&lt;211&gt; 212

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:449941.2.orf1:2000SEP08

&lt;400&gt; 490

```

Ala Ala Ala Arg Ala Asp Lys Glu Gln Ala Glu Glu Glu Leu Cys
 1          5          10          15
Gln Ala Thr Met Gly Ile Ser Arg Asp Ser Met His Lys Arg Arg
          20          25          30
Ala Thr Gly Gly Lys Gln Lys Ala Trp Arg Lys Lys Arg Lys Tyr
          35          40          45
Glu Leu Gly Arg Gln Pro Ala Asn Thr Lys Leu Ser Ser Asn Lys
          50          55          60
Thr Val Arg Arg Val Arg Val Arg Gly Gly Asn Val Lys Trp Arg
          65          70          75
Ala Leu Arg Leu Asp Thr Gly Asn Tyr Ser Trp Gly Ser Glu Ala
          80          85          90
Val Thr Arg Lys Thr Arg Ile Leu Asp Val Val Tyr Asn Ala Ser
          95          100          105
Asn Asn Glu Leu Val Arg Thr Gln Thr Leu Val Lys Ser Ala Ile
          110          115          120
Val Gln Val Asp Ala Ala Pro Phe Lys Gln Trp Tyr Leu Thr His
          125          130          135
Tyr Gly Val Asp Ile Gly Arg Lys Lys Lys Thr Pro Ala Ala Lys
          140          145          150
Lys Asp Asn Ala Glu Gly Gln Glu Val Glu Ala Ala Ala Glu Glu

```

	155		160		165
Thr Lys Lys Ser	Asn His Val Thr Arg Lys	Leu Glu Lys Arg Lys			
	170		175		180
Glu Gly Arg Thr	Leu Asp Pro His Ile Glu Glu Gln Tyr Trp Gln				
	185		190		195
Trp Thr Val Ala	Gly Met His Phe Phe Pro Pro Trp Thr Val Trp				
	200		205		210
Pro Ser					

&lt;210&gt; 491

&lt;211&gt; 217

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450229.1.orf1:2000SEP08

&lt;400&gt; 491

Leu Ala Pro Ala Pro	Pro Leu Arg Phe Pro His Ser Leu Leu Gln	
1	5	10
Leu Arg Val Arg Arg	Arg Arg Glu Gly Lys Met Tyr Thr Ala Arg	
	20	25
Lys Lys Ile Gln Lys	Glu Lys Gly Leu Glu Pro Ser Glu Phe Glu	
	35	40
Asp Ser Val Ala Gln	Ala Phe Phe Asp Leu Glu Asn Gly Asn Gln	
	50	55
Glu Leu Lys Ser Asp	Leu Lys Asp Leu Tyr Ile Asn Asn Ala Ile	
	65	70
Gln Met Asp Val Thr	Gly Ser Arg Lys Ala Val Val Ile His Val	
	80	85
Pro Tyr Arg Leu Arg	Lys Ala Phe Arg Lys Ile His Val Arg Leu	
	95	100
Val Arg Glu Leu Glu	Lys Lys Phe Ser Gly Lys Asp Val Val Ile	
	110	115
Val Ala Thr Arg Arg	Ile Val Arg Pro Pro Lys Lys Gly Ser Ala	
	125	130
Val Leu Arg Pro Arg	Thr Arg Thr Leu Thr Ala Val His Asp Gly	
	140	145
Ile Leu Glu Asp Val	Val Tyr Pro Ala Glu Ile Val Gly Lys Arg	
	155	160
Val Arg Tyr Arg Leu	Asp Gly Ser Lys Ile Ile Lys Ile Phe Leu	
	170	175
Asp Pro Lys Glu Arg	Asn Asn Thr Glu Tyr Lys Leu Glu Thr Cys	
	185	190
Thr Ala Val Tyr Arg	Arg Leu Cys Gly Lys Asp Val Val Phe Glu	
	200	205
Tyr Pro Met Thr Glu	Asn Ala	
	215	

&lt;210&gt; 492

&lt;211&gt; 148

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450399.3.orf3:2000SEP08

&lt;400&gt; 492

Pro Pro Leu Leu Ser	Ala Ser Lys Met Ser Lys Arg Gly Arg Gly
1	5
	10
	15

Gly Ser Ala Gly Asn Lys Phe Arg Met Ser Leu Gly Leu Pro Val  
 20 25 30  
 Ala Ala Thr Val Asn Cys Ala Asp Asn Thr Gly Ala Lys Asn Leu  
 35 40 45  
 Tyr Ile Ile Ser Val Lys Gly Ile Lys Gly Arg Leu Asn Arg Leu  
 50 55 60  
 Pro Ser Ala Cys Val Gly Asp Met Val Met Ala Thr Val Lys Lys  
 65 70 75  
 Gly Lys Pro Asp Leu Arg Lys Lys Val Met Pro Ala Val Ile Val  
 80 85 90  
 Arg Gln Arg Lys Pro Trp Arg Arg Lys Asp Gly Val Tyr Met Tyr  
 95 100 105  
 Phe Glu Asp Asn Ala Gly Val Ile Val Asn Pro Lys Gly Glu Met  
 110 115 120  
 Lys Gly Ser Ala Ile Thr Gly Pro Ile Gly Lys Glu Cys Ala Asp  
 125 130 135  
 Leu Trp Pro Arg Ile Ala Ser Ala Ala Asn Ala Ile Val  
 140 145

&lt;210&gt; 493

&lt;211&gt; 158

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:455771.1.orf3:2000SEP08

&lt;400&gt; 493

Ala Ser Cys Ser Arg Arg Arg Glu Ala Leu Gln Arg Thr Ser Val  
 1 5 10 15  
 Asn Met Gly Lys Thr Arg Gly Met Gly Ala Gly Arg Lys Leu Lys  
 20 25 30  
 Thr His Arg Arg Asn Gln Arg Trp Ala Asp Lys Ala Tyr Lys Lys  
 35 40 45  
 Ser His Leu Gly Asn Glu Trp Lys Lys Pro Phe Ala Gly Ser Ser  
 50 55 60  
 His Ala Lys Gly Ile Val Leu Glu Lys Ile Gly Ile Glu Ala Lys  
 65 70 75  
 Gln Pro Asn Ser Ala Ile Arg Lys Cys Ala Arg Val Gln Leu Val  
 80 85 90  
 Lys Asn Gly Lys Lys Ile Ala Ala Phe Val Pro Asn Asp Gly Cys  
 95 100 105  
 Leu Asn Tyr Ile Glu Glu Asn Asp Glu Val Leu Ile Ala Gly Phe  
 110 115 120  
 Gly Arg Lys Gly His Ala Val Gly Asp Ile Pro Gly Val Arg Phe  
 125 130 135  
 Lys Val Val Lys Val Ser Gly Val Ser Leu Leu Ala Leu Phe Lys  
 140 145 150  
 Glu Lys Lys Glu Lys Pro Arg Ser  
 155

&lt;210&gt; 494

&lt;211&gt; 188

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:720459.1.orf1:2000SEP08

&lt;400&gt; 494

Val Leu Ala Thr Val Thr Lys Thr Val Gly Gly Asp Lys Asn Gly

1	5	10	15
Gly Thr Arg Val Val Lys Leu Arg Lys Met Pro Arg Tyr Tyr Pro	20	25	30
Thr Glu Asp Val Pro Arg Lys Leu Leu Ser His Gly Lys Lys Pro	35	40	45
Phe Ser Gln His Val Arg Arg Leu Arg Ser Ser Ile Thr Pro Gly	50	55	60
Thr Val Leu Ile Ile Leu Thr Gly Arg His Arg Gly Lys Arg Val	65	70	75
Val Phe Leu Lys Gln Leu Gly Ser Gly Leu Leu Leu Val Thr Gly	80	85	90
Pro Leu Ala Leu Asn Arg Val Pro Leu Arg Arg Thr His Gln Lys	95	100	105
Phe Val Ile Ala Thr Ser Thr Lys Val Asp Ile Ser Lys Val Lys	110	115	120
Ile Pro Lys His Leu Thr Asp Ala Tyr Phe Lys Lys Lys Pro Leu	125	130	135
Arg Lys Pro Arg His Gln Glu Gly Glu Ile Phe Asp Thr Glu Lys	140	145	150
Glu Lys Tyr Glu Ile Thr Glu Gln Arg Lys Ala Asp Gln Lys Ala	155	160	165
Val Asp Ser Gln Ile Leu Pro Lys Ile Lys Ala Val Pro Gln Leu	170	175	180
Gln Gly Tyr Leu Arg Ser Gln Phe	185		

&lt;210&gt; 495

&lt;211&gt; 144

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:723156.1.orf3:2000SEP08

&lt;400&gt; 495

Arg Arg Asn Leu Pro Arg Ala Arg Ala Asn Lys Gly Gly Arg Gly	1	5	10	15
Ala Gly Gly Gly Ala Glu Met Ala Glu Lys Lys Arg Gly Ala Gly	20	25	30	
Thr Arg Lys Asp Glu Val Val Thr Arg Glu Tyr Thr Ile Asn Leu	35	40	45	
His Lys Arg Leu His Gly Cys Thr Phe Lys Lys Lys Ala Pro Asn	50	55	60	
Ala Ile Lys Glu Ile Arg Lys Phe Ala Gln Lys Ala Met Gly Thr	65	70	75	
Thr Asp Val Arg Ile Asp Val Lys Leu Asn Lys His Ile Trp Ser	80	85	90	
Ser Gly Ile Arg Ser Val Pro Arg Arg Val Arg Val Arg Ile Ala	95	100	105	
Arg Asn Arg Asn Asp Glu Glu Asp Ala Lys Glu Glu Leu Tyr Ser	110	115	120	
Leu Val Thr Val Ala Glu Ile Pro Pro Glu Gly Leu Lys Gly Leu	125	130	135	
Gly Thr Lys Val Val Glu Asp Glu Asp	140			

&lt;210&gt; 496

&lt;211&gt; 159

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:728055.1.orf3:2000SEP08

&lt;400&gt; 496

```

Arg Gly Ala Gly Glu Glu Leu Val Phe Ile Leu Val Ala Met Pro
 1          5          10          15
Ser Lys Gly Pro Leu Gln Ser Val Gln Val Phe Gly Arg Lys Lys
          20          25          30
Thr Ala Thr Ala Val Ala His Cys Lys Arg Gly Asn Gly Leu Ile
          35          40          45
Lys Val Asn Gly Arg Pro Leu Glu Met Ile Glu Pro Arg Thr Leu
          50          55          60
Gln Tyr Lys Leu Leu Glu Pro Val Leu Leu Gly Lys Glu Arg
          65          70          75
Phe Ala Gly Val Asp Ile Arg Val Arg Val Lys Gly Gly Gly His
          80          85          90
Val Ala Gln Ile Tyr Ala Ile Arg Gln Ser Ile Ser Lys Ala Leu
          95          100          105
Val Ala Tyr Tyr Gln Lys Tyr Val Asp Glu Ala Ser Lys Lys Glu
          110          115          120
Ile Lys Asp Ile Leu Ile Gln Tyr Asp Arg Thr Leu Leu Val Ala
          125          130          135
Asp Pro Cys Arg Cys Glu Ser Lys Lys Phe Gly Gly Pro Gly Ala
          140          145          150
Arg Ala Arg Tyr Gln Lys Ser Tyr Arg
          155

```

&lt;210&gt; 497

&lt;211&gt; 147

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1020789.1.orf1:2000SEP08

&lt;400&gt; 497

```

Gly Leu His Ala Ala Ala Cys Ala Ala Met Ser Leu Val Ile
 1          5          10          15
Pro Glu Lys Phe Gln His Ile Leu Arg Val Leu Asn Thr Asn Ile
          20          25          30
Asp Gly Arg Arg Lys Ile Ala Phe Ala Ile Thr Ala Ile Lys Gly
          35          40          45
Val Gly Arg Arg Tyr Ala His Val Val Leu Arg Lys Ala Asp Ile
          50          55          60
Asp Leu Thr Lys Arg Ala Gly Glu Leu Thr Glu Asp Glu Val Glu
          65          70          75
Arg Val Ile Thr Ile Met Gln Asn Pro Arg Gln Tyr Lys Ile Pro
          80          85          90
Asp Trp Phe Leu Asn Arg Gln Lys Asp Val Lys Asp Gly Lys Tyr
          95          100          105
Ser Gln Val Leu Ala Asn Gly Leu Asp Asn Lys Leu Arg Glu Asp
          110          115          120
Leu Glu Arg Leu Lys Lys Ile Arg Ala His Arg Gly Leu Arg His
          125          130          135
Phe Trp Gly Leu Arg Val Arg Gly Gln His Thr Lys
          140          145

```

&lt;210&gt; 498

&lt;211&gt; 130

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens



&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1071728.1.orf1:2000SEP08

&lt;400&gt; 498

Gly	Thr	Ser	Arg	Ser	Gly	Ser	Tyr	Ala	Ser	Ala	Met	Ala	Phe	Lys
1				5					10					15
Asp	Thr	Gly	Lys	Thr	Pro	Val	Glu	Pro	Glu	Val	Ala	Ile	His	Arg
				20					25					30
Ile	Arg	Ile	Thr	Leu	Thr	Ser	Arg	Asn	Val	Lys	Ser	Leu	Glu	Lys
				35					40					45
Val	Cys	Ala	Asp	Leu	Ile	Arg	Gly	Ala	Lys	Glu	Lys	Asn	Leu	Lys
				50					55					60
Val	Lys	Gly	Pro	Val	Arg	Met	Pro	Thr	Lys	Thr	Leu	Arg	Ile	Thr
				65					70					75
Thr	Arg	Lys	Thr	Pro	Cys	Gly	Glu	Gly	Ser	Lys	Thr	Trp	Asp	Arg
				80					85					90
Phe	Gln	Met	Arg	Ile	His	Lys	Arg	Leu	Ile	Asp	Leu	His	Ser	Pro
				95					100					105
Ser	Glu	Ile	Val	Lys	Gln	Ile	Thr	Ser	Ile	Ser	Ile	Glu	Pro	Gly
				110					115					120
Val	Glu	Val	Glu	Val	Thr	Ile	Ala	Asp	Ala					
				125					130					

&lt;210&gt; 499

&lt;211&gt; 149

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1084329.1.orf2:2000SEP08

&lt;400&gt; 499

Glu	Ala	Pro	Ala	Pro	Pro	Lys	Ala	Glu	Ala	Lys	Ala	Lys	Ala	Leu
1				5					10					15
Lys	Ala	Lys	Lys	Ala	Val	Leu	Lys	Gly	Val	His	Ser	His	Lys	Lys
				20					25					30
Lys	Lys	Ile	Arg	Thr	Ser	Pro	Thr	Phe	Arg	Arg	Pro	Lys	Thr	Leu
				35					40					45
Arg	Leu	Arg	Arg	Gln	Pro	Lys	Tyr	Pro	Arg	Lys	Ser	Ala	Pro	Arg
				50					55					60
Arg	Asn	Lys	Leu	Asp	His	Tyr	Ala	Ile	Ile	Lys	Phe	Pro	Leu	Thr
				65					70					75
Thr	Glu	Ser	Ala	Met	Lys	Lys	Ile	Glu	Asp	Asn	Asn	Thr	Leu	Val
				80					85					90
Phe	Ile	Val	Asp	Val	Lys	Ala	Asn	Lys	His	Gln	Ile	Lys	Gln	Ala
				95					100					105
Val	Lys	Lys	Leu	Tyr	Asp	Ile	Asp	Val	Ala	Lys	Val	Asn	Thr	Leu
				110					115					120
Ile	Arg	Pro	Asp	Gly	Glu	Lys	Lys	Ala	Tyr	Val	Arg	Leu	Ala	Pro
				125					130					135
Asp	Tyr	Asp	Ala	Leu	Asp	Val	Ala	Asn	Lys	Ile	Gly	Ile	Ile	
				140					145					

&lt;210&gt; 500

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:246422.1.orf1:2000SEP08

&lt;400&gt; 500

```

Lys Lys Arg Lys Gln Val Pro Lys Phe Thr Leu Asp Arg Thr His
 1              5              10              15
Pro Val Glu Asp Gly Ile Met Asp Ala Ala Asn Phe Glu Gln Phe
              20              25              30
Phe Gln Glu Arg Ile Lys Met Asn Gly Lys Ala Gly Asn Phe Gly
              35              40              45
Gly Gly Val Val Thr Ile Glu Gly Ser Lys Ser Lys Thr Ser Val
              50              55              60
Thr Ser Lys Leu Pro Phe Ser Asn Arg Tyr Leu Lys Tyr Leu Thr
              65              70              75
Lys Lys Tyr Leu Lys Lys Asn Asn Leu His Asp Trp Leu Arg Val
              80              85              90
Val Ala Asn Ser Lys Gln Ser Tyr Glu Leu Arg Tyr Phe Gln Ile
              95              100             105
Asn Gln Asp Glu Glu Glu Glu Asn Glu Asp
              110             115

```

&lt;210&gt; 501

&lt;211&gt; 178

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1086066.1.orf2:2000SEP08

&lt;400&gt; 501

```

Pro Glu Leu Arg Val Val Arg Ala Asp Arg Leu Arg Ser Pro Val
 1              5              10              15
Cys Ala Leu Arg Met Thr Glu Trp Glu Thr Ala Thr Pro Ala Val
              20              25              30
Ala Glu Thr Pro Asp Ile Lys Leu Phe Gly Lys Trp Ser Thr Asp
              35              40              45
Asp Val Gln Ile Asn Asp Ile Ser Leu Gln Asp Tyr Ile Ala Val
              50              55              60
Lys Glu Lys Tyr Ala Lys Tyr Leu Pro His Ser Ala Gly Arg Tyr
              65              70              75
Ala Ala Lys Arg Phe Arg Lys Ala Gln Cys Pro Ile Val Glu Arg
              80              85              90
Leu Thr Asn Ser Met Met Met His Gly Arg Asn Asn Gly Lys Lys
              95              100             105
Leu Met Thr Val Arg Ile Val Lys His Ala Phe Glu Ile Ile His
              110             115             120
Leu Leu Thr Gly Glu Asn Pro Leu Gln Val Leu Val Asn Ala Ile
              125             130             135
Ile Asn Ser Gly Pro Arg Glu Asp Ser Thr Arg Ile Gly Arg Ala
              140             145             150
Gly Thr Val Arg Arg Gln Ala Val Asp Val Ser Pro Leu Arg Arg
              155             160             165
Val Asn Gln Ala Ile Trp Leu Leu Cys Ile Gly Gly Ser
              170             175

```

&lt;210&gt; 502

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:223142.1.orf3:2000SEP08

&lt;400&gt; 502

```

Arg Cys Ser Gly His His Arg Tyr Ala Tyr Met Lys Ser His Gln
 1      5      10
Val Ser Glu Ile Phe Ser Phe Thr Glu Ile Met Cys Thr Ile Ser
 20      25      30
Val Leu Glu Ser Gly Val Gly Arg Cys Thr Gln Asp Arg His Trp
 35      40      45
Gly Trp Thr His His Gln Trp Pro Arg Lys Gly Thr Glu Ile Cys
 50      55      60
Leu Gln Val Gln Ser Tyr Ala Glu Leu Lys Gly Ile Asp Val Asp
 65      70      75
Ser Leu Val Ile Glu His Ile Gln Gly Lys Gly Thr His Asn Val
 80      85      90
Pro Pro Asp Leu Gln Asn Ser Trp Ala Asp Glu Pro Ile His Lys
 95     100     105
Leu Pro Cys His Ile Gln Met Met Leu Ser Glu Lys Lys His Leu
110     115     120
Val Pro Lys Ala Glu Lys Glu Asp Ala Arg Lys Lys Lys Ile Pro
125     130     135
Gln Lys Lys His Lys Leu Lys Arg Gln Thr Asn Ser Ala Lys Arg
140     145     150
Lys Cys Lys

```

&lt;210&gt; 503

&lt;211&gt; 155

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:885368.1.orf2:2000SEP08

&lt;400&gt; 503

```

Pro Leu Leu Leu Gln Ser Ser Ser Pro Pro Ala Thr Ala Thr Leu
 1      5      10
Asp His Ala Val Arg Arg Gly Lys Val Ser Met Val Arg Val Ser
 20      25      30
Val Leu Asn Asp Ala Leu Lys Ser Met Tyr Asn Ala Glu Lys Ile
 35      40      45
Gly Lys Arg Gln Val Met Ile Arg Pro Ser Ser Lys Val Ile Ile
 50      55      60
Lys Phe Leu Thr Val Met Gln Arg His Gly Tyr Ile Gly Glu Phe
 65      70      75
Glu Tyr Val Asp Asp His Arg Ser Gly Lys Ile Val Val Glu Leu
 80      85      90
Asn Gly Arg Leu Asn Lys Cys Gly Val Ile Ser Pro Arg Phe Asp
 95     100     105
Ile Gly Val Lys Asp Ile Glu Gly Trp Thr Ala Arg Leu Leu Pro
110     115     120
Ser Arg Gln Phe Gly Tyr Ile Val Leu Thr Thr Ser Ala Gly Ile
125     130     135
Met Asp His Glu Glu Ala Arg Arg Lys Ser Val Gly Gly Lys Val
140     145     150
Leu Gly Phe Phe Tyr
155

```

&lt;210&gt; 504

&lt;211&gt; 98

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: LI:481782.1.orf2:2000SEP08

<400> 504

Cys	Asn	Ala	Met	Tyr	Leu	Leu	Met	Ala	Arg	Ser	Gly	Gln	Thr	Arg	
1				5					10					15	
Pro	Thr	Leu	Leu	Gly	Phe	Met	Asp	Val	Ile	Ser	Ile	Pro	Lys	Thr	
				20					25					30	
Asn	Glu	Asn	Tyr	Arg	Leu	Leu	Tyr	Asp	Thr	Lys	Gly	Arg	Phe	Arg	
				35					40					45	
Leu	His	Pro	Ile	Arg	Asp	Glu	Asp	Ala	Lys	Phe	Lys	Leu	Cys	Lys	
				50					55					60	
Val	Arg	Ser	Val	Gln	Leu	Gly	Gln	Lys	Gly	Ile	Pro	Tyr	Leu	Asn	
				65					70					75	
Thr	Tyr	Asp	Gly	Arg	Thr	Ile	Arg	Tyr	Pro	Asp	Pro	Leu	Ile	Lys	
				80					85					90	
Ala	Asn	Asp	Thr	Ile	Lys	Ile	Asp								
				95											

<210> 505

<211> 132

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:1093813.1.orf1:2000SEP08

<400> 505

Ala	Ala	Ser	Ser	Ala	Asn	Ser	Ala	Ala	Met	Ala	Lys	Ile	Lys	Ala	
1				5					10					15	
Arg	Asp	Leu	Arg	Gly	Lys	Lys	Lys	Glu	Glu	Leu	Leu	Lys	Gln	Leu	
				20					25					30	
Asp	Asp	Leu	Lys	Val	Glu	Leu	Ser	Gln	Leu	Arg	Val	Ala	Lys	Val	
				35					40					45	
Thr	Gly	Gly	Ala	Ala	Ser	Lys	Leu	Ser	Lys	Ile	Arg	Val	Val	Arg	
				50					55					60	
Lys	Ser	Ile	Ala	Arg	Val	Leu	Thr	Val	Ile	Asn	Gln	Thr	Gln	Lys	
				65					70					75	
Glu	Asn	Leu	Arg	Lys	Phe	Tyr	Lys	Gly	Lys	Lys	Tyr	Lys	Pro	Leu	
				80					85					90	
Asp	Leu	Arg	Pro	Lys	Lys	Thr	Arg	Ala	Met	Arg	Arg	Arg	Leu	Thr	
				95					100					105	
Lys	His	Glu	Glu	Lys	Leu	Lys	Thr	Lys	Lys	Gln	Gln	Arg	Lys	Glu	
				110					115					120	
Arg	Leu	Tyr	Pro	Leu	Arg	Lys	Tyr	Ala	Val	Lys	Ala				
				125					130						

<210> 506

<211> 163

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:449413.2.orf3:2000SEP08

<400> 506

Pro	Arg	Arg	Phe	Arg	Leu	Pro	Gln	Arg	Arg	Arg	Pro	Ser	Gln	Pro	
1				5					10					15	
Val	Pro	Ser	Ser	Ala	Thr	Met	Gly	Lys	Thr	Arg	Gly	Met	Gly	Ala	
				20					25					30	
Gly	Arg	Lys	Leu	Lys	Thr	His	Arg	Arg	Asn	Gln	Arg	Trp	Ala	Asp	
				35					40					45	

```

Lys Ala Tyr Lys Lys Ser His Leu Gly Asn Glu Trp Lys Lys Pro
      50      55      60
Phe Ala Gly Ser Ser His Ala Lys Gly Ile Val Leu Glu Lys Ile
      65      70      75
Gly Ile Glu Ala Lys Gln Pro Asn Ser Ala Ile Arg Lys Cys Ala
      80      85      90
Arg Val Gln Leu Val Lys Asn Gly Lys Lys Ile Ala Ala Phe Val
      95     100     105
Pro Asn Asp Gly Cys Leu Asn Tyr Ile Glu Glu Asn Asp Glu Val
     110     115     120
Leu Ile Ala Gly Phe Gly Arg Lys Gly His Ala Val Gly Asp Ile
     125     130     135
Pro Gly Val Arg Phe Lys Val Val Lys Val Ser Gly Val Ser Leu
     140     145     150
Leu Ala Leu Phe Lys Glu Lys Lys Glu Lys Pro Arg Ser
     155     160

```

&lt;210&gt; 507

&lt;211&gt; 119

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450105.1.orf1:2000SEP08

&lt;400&gt; 507

```

Gln Ser Leu Gln Glu Glu Pro Ser Arg Leu Thr Ser Gly Arg Asn
  1      5      10      15
Pro Leu Leu Gly His Leu Thr Gln Arg Glu Ser Ser Leu Arg Arg
     20     25     30
Ser Ala Leu Arg Leu Ser Ser Leu Thr Leu Leu Ser Val Ser Ala
     35     40     45
Leu Val Leu Gln Leu Val Lys Asn Gly Lys Lys Ile Ala Ala Phe
     50     55     60
Val Pro Asn Asp Gly Cys Leu Asn Tyr Ile Glu Glu Asn Asp Glu
     65     70     75
Val Leu Ile Ala Gly Phe Gly Arg Lys Gly His Ala Val Gly Asp
     80     85     90
Ile Pro Gly Val Arg Phe Lys Val Val Lys Val Ser Gly Val Ser
     95    100    105
Leu Leu Ala Leu Phe Lys Glu Lys Lys Glu Lys Pro Arg Ser
    110    115

```

&lt;210&gt; 508

&lt;211&gt; 119

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:450105.1.orf3:2000SEP08

&lt;400&gt; 508

```

Arg Arg Gln Leu Gly Arg Arg Thr Pro Ala Thr Glu Leu Gly Gly
  1      5      10      15
Ile Thr Phe Ala Asp Pro Thr Trp Val Arg His Val Val Trp Glu
     20     25     30
Leu Gly Ala Ser Ser Arg Pro Thr Glu Gly Thr Arg Gly Gly Leu
     35     40     45
Thr Lys Pro Thr Arg Arg Ala Ile Ser Ala Asn Glu Trp Lys Lys
     50     55     60
Pro Phe Ala Gly Ser Ser His Ala Lys Gly Ile Val Leu Glu Lys

```

	65		70		75
Ile Gly Ile Glu Ala	Lys Gln Pro Asn Ser	Ala Ile Arg Lys Cys			
	80		85		90
Ala Arg Ala Pro Ala	Gly Glu Glu Arg	Glu Glu Asp Cys Cys	Leu		
	95		100		105
Cys Ala Lys Arg Trp	Leu Leu Glu Leu	His Arg Gly Glu Arg			
	110		115		

&lt;210&gt; 509

&lt;211&gt; 107

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:814285.1.orf1:2000SEP08

&lt;400&gt; 509

Arg Ser Ser Asp Leu	Asp Glu Ile Ser	Leu Ala Pro Lys	Asp Ala
1	5	10	15
Ser Lys Gly Gln Gly	Gly Ser Leu Ser	Tyr Glu Lys Ala	Lys His
	20	25	30
Val Ser Gln Gly Asn	Ile Cys Arg Cys	Thr Glu Leu Lys	Phe Lys
	35	40	45
Ile Ser Arg Ile Ala	Arg Lys Ala Gly	Asn Phe Tyr Val	Ser Ala
	50	55	60
Glu Pro Lys Leu Ala	Phe Val Ile Arg	Ile Gly Gly Tyr	His Ser
	65	70	75
Gly Trp Ser Pro Lys	Gly Leu Lys Arg	Cys Cys Lys Leu	Leu Cys
	80	85	90
Leu His Gln Phe Phe	His Glu Gln Leu	Cys Glu Ser Ser	Ala Gly
	95	100	105
Leu Gln			

&lt;210&gt; 510

&lt;211&gt; 102

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1142855.1.orf2:2000SEP08

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 36

&lt;223&gt; unknown or other

&lt;400&gt; 510

Ser Lys Gln Pro Lys	Ala His Val Ile	Ile Arg Leu Leu	Arg Phe
1	5	10	15
Phe Phe Val Cys Phe	His Phe Leu Leu	Ser Trp Ser Ser	Ser Ser
	20	25	30
Gly Val Gly Arg Thr	Xaa Cys Trp Cys	Ser Thr Ser Cys	Trp Ser
	35	40	45
Arg Ser Thr Ser Pro	Tyr Ile Val Asp	Glu Ala Pro Asn	Val Asp
	50	55	60
Ile Gly Gln Gly His	Cys Lys Gln Ala	Gly Pro Lys Arg	Phe Asn
	65	70	75
Asn Leu His Gln Leu	His Leu Thr Arg	Ala Leu Ile Leu	Ser Ser
	80	85	90
Val Met Val Thr Ser	Ile Val Met Gln	Ser Glu Gly	

95

100

<210> 511  
 <211> 82  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:817330.1.orf3:2000SEP08

<400> 511  
 Gly Thr Leu Pro Gln Phe Leu Gly Gln Pro Met Asp Leu Met Thr  
 1 5 10 15  
 Ala Leu Gln Leu Val Met Lys Lys Ser Ser Ala His Asp Gly Leu  
 20 25 30  
 Val Lys Gly Leu Arg Val Gly Cys Gln Gly Leu Ser Arg Ser Thr  
 35 40 45  
 Pro Leu Ser Leu Cys Val Leu Ala Glu Asp Cys Asp Gln Pro Asp  
 50 55 60  
 Tyr Val Lys Leu Val Lys Ala Leu Cys Ala Glu His Asn Val His  
 65 70 75  
 Leu Val Thr Val Pro Ala Asp  
 80

<210> 512  
 <211> 166  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:817845.1.orf2:2000SEP08

<400> 512  
 Ser Glu Gln Pro His Arg Thr Leu Pro Ala Phe Leu Ser Ser Pro  
 1 5 10 15  
 Gly Gly Val Arg Gly Arg Met Ala Glu Ser Thr Ala Arg Thr Val  
 20 25 30  
 Lys Asp Val Asn Pro His Glu Phe Val Lys Ala Tyr Ser Ala His  
 35 40 45  
 Leu Lys Arg Ser Gly Lys Met Glu Leu Pro Glu Trp Val Asp Ile  
 50 55 60  
 Val Lys Thr Ala Arg Phe Lys Glu Leu Pro Pro Tyr Asp Pro Asp  
 65 70 75  
 Trp Tyr Tyr Ile Arg Ala Ala Ser Ile Ala Arg Lys Ile Tyr Leu  
 80 85 90  
 Arg Gln Gly Ile Gly Val Gly Gly Phe Gln Lys Ile Tyr Gly Gly  
 95 100 105  
 Arg Gln Arg Asn Gly Ser Arg Pro Pro His Phe Cys Lys Ser Ser  
 110 115 120  
 Gly Ala Val Ala Arg Asn Ile Leu Gln Gln Leu Gln Ile Met Gly  
 125 130 135  
 Ile Ile Asp Val Asp Pro Lys Gly Gly Arg Leu Ile Thr Asn Gln  
 140 145 150  
 Gly Arg Arg Asp Leu Asp Gln Val Ala Gly Arg Val Ala Val Glu  
 155 160 165  
 Ala

<210> 513  
 <211> 125  
 <212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:460809.1.orf1:2000SEP08

<400> 513

Leu	Gln	Arg	Arg	Val	Ala	Arg	Arg	Lys	Val	Val	Pro	Pro	Leu	Val
1				5					10					15
Arg	Arg	Leu	Pro	Gln	Asn	Thr	Pro	Gly	Ala	Trp	Val	Glu	Trp	Ala
				20					25					30
Ser	Arg	Ser	Ala	Pro	Arg	Ala	His	Arg	Glu	Ile	Gln	Lys	Phe	Ala
				35					40					45
Met	Lys	Glu	Met	Gly	Thr	Pro	Asn	Leu	His	Ile	Asp	Val	Arg	Leu
				50					55					60
Asn	Lys	Ala	Leu	Trp	Ala	Lys	Gly	Ile	Arg	Asn	Val	Pro	Tyr	His
				65					70					75
Ile	His	Met	Lys	Leu	Pro	Arg	Lys	Leu	Asn	Glu	Asp	Glu	Asp	Ser
				80					85					90
Pro	Asp	Lys	Leu	Tyr	Ala	Leu	Val	Pro	Thr	Tyr	Thr	Cys	Tyr	His
				95					100					105
Phe	His	Lys	Ser	Ile	Asp	Arg	Gln	Cys	Gly	Arg	Glu	Leu	Thr	Thr
				110					115					120
Asp	Gly	Ser	Ile	His										
				125										

<210> 514

<211> 91

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:815874.1.orf3:2000SEP08

<400> 514

Gln	Phe	Thr	Glu	Phe	Glu	Ser	Leu	Ser	Met	Ser	His	Ser	Ile	Lys
1				5					10					15
Ser	Leu	Glu	Lys	Val	Cys	Ala	Asp	Leu	Ile	Arg	Gly	Ala	Lys	Lys
				20					25					30
Lys	Asn	Leu	Lys	Val	Lys	Gly	Pro	Val	Gln	Met	Pro	Thr	Lys	Thr
				35					40					45
Leu	Arg	Ile	Ala	Thr	Arg	Lys	Thr	Pro	Phe	Gly	Asp	Gly	Ser	Lys
				50					55					60
Thr	Trp	Asp	His	Phe	His	Met	Arg	Ile	His	Lys	Gln	Leu	Ile	Asp
				65					70					75
Leu	His	Ile	Leu	Ser	Glu	Ile	Val	Lys	Gln	Ile	Ser	Phe	His	Gln
				80					85					90
Tyr														

<210> 515

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:255713.1.orf2:2000SEP08

<400> 515

Tyr	Gln	Ala	Gln	Gln	Ser	Cys	Leu	Asp	Gln	Arg	Asn	Lys	Glu	Arg
1				5					10					15



Ser Ile Leu Cys Pro Cys Ala Val Asp Gly Pro Glu Asn Ile Met  
 20 25 30  
 Arg Leu Glu Ile His Gln Ile Ser Ser Ile Leu Trp Leu Pro Val  
 35 40 45  
 Tyr Pro Leu Pro Leu Ser Glu Thr Asp Ser Gln Thr Leu Trp Met  
 50 55 60  
 Arg Thr Asn Cys

<210> 516  
 <211> 100  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:035973.1.orf1:2000SEP08

<400> 516  
 Gly Cys Leu Ala Gly Ile Arg Lys Asp Asn Lys Met Lys Gly Thr  
 1 5 10 15  
 Ser Pro Phe Gly Lys Cys Arg Asp Met Ile His Lys Leu Cys Cys  
 20 25 30  
 Leu Cys Gly Ser Lys Ala Tyr His Leu Gln Lys Leu Thr Cys Gly  
 35 40 45  
 Lys Cys Gly Tyr Pro Ala Lys Arg Ile Val Glu Ser Ile Tyr Gly  
 50 55 60  
 Ser Ala Lys Ala Arg Ile Asp Leu Tyr Thr Ser Gly Thr Gly Arg  
 65 70 75  
 Met Arg His Leu Lys Asn Cys Arg Ser Ala Asp Ser Gly Thr Asp  
 80 85 90  
 Ser Val Lys Glu His His Leu Asn Pro Lys  
 95 100

<210> 517  
 <211> 83  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:1138110.1.orf3:2000SEP08

<400> 517  
 Arg Lys Lys Trp Thr Leu Cys Tyr Val Phe Pro Tyr Phe Val Leu  
 1 5 10 15  
 Ser Val Tyr Ser Tyr Ile Asp Ser Val Val Ala Met His Asn Phe  
 20 25 30  
 Pro Thr Asn Phe Ile Leu Thr Cys Arg Glu Ser Val Asp Lys Ile  
 35 40 45  
 Phe Cys Asn Lys Val Leu Phe Ala Asn Met Tyr Phe Ile Phe Thr  
 50 55 60  
 Val Tyr Ser Ile Phe Leu Ile Pro Tyr Lys Phe Leu Gln Glu Ser  
 65 70 75  
 Phe Arg Phe Ser Ile Gln Asn Gly  
 80

<210> 518  
 <211> 188  
 <212> PRT  
 <213> Homo sapiens

<220>

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2049074.1.orf2:2000SEP08

&lt;400&gt; 518

```

Phe Ser Cys Thr Gly Arg Gln Pro Trp Arg Pro Ala Gly Ile Gly
 1          5          10          15
Met Ile Leu Lys Pro His Phe Gln Lys Glu Trp Gln Ala Gly Ser
 20          25          30
Gly Thr Arg Gly Ser Thr Ser Arg Pro Ala Arg Ser Ala Asp Ala
 35          40          45
Arg Pro Gly Arg Arg Lys Arg Ala Ala Ser Pro Leu Ala Pro Arg
 50          55          60
Ser Gly Pro Ile Arg Pro Ile Val Arg Cys Pro Thr Val Arg Tyr
 65          70          75
His Ile Gln Gly Pro Gly Trp Gln Gly Leu Gln Pro Gly Gly Ala
 80          85          90
Gln Gly Gly Trp Tyr Pro Gln Glu Asn Gly Thr His His Arg His
 95          100         105
Leu Arg Gly Pro Lys Glu Ala Lys Gln Ile His Gly Val Thr Ala
 110         115         120
Gly Gln Arg Ala Ala Pro Glu Gly Val Pro Leu Gln Ala His Thr
 125         130         135
Phe Pro Pro Gly Ser Pro Ser Ala Pro Glu Val Arg Glu Thr Val
 140         145         150
Leu Leu Lys Asn Leu Asn Trp Gln Arg Ser Leu Thr Met Val Pro
 155         160         165
Val Met Leu Tyr Pro Glu Cys Val Gln Lys Gly Glu Gly Gln Ser
 170         175         180
His His Gly Arg Gly Glu Glu Leu
 185

```

&lt;210&gt; 519

&lt;211&gt; 143

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1092460.1.orf1:2000SEP08

&lt;400&gt; 519

```

Ala Ala Ala Ala Arg Gly Cys Ser Gly Arg Arg Ala Ala Glu Thr
 1          5          10          15
Met Ala Gly Gly Lys Ala Gly Lys Asp Ser Gly Lys Ala Lys Ala
 20          25          30
Lys Ala Val Ser Arg Ser Gln Arg Ala Gly Leu Gln Phe Pro Val
 35          40          45
Gly Arg Ile His Arg His Leu Lys Thr Arg Thr Thr Ser His Gly
 50          55          60
Arg Val Gly Ala Thr Ala Ala Val Tyr Ser Ala Ala Ile Leu Glu
 65          70          75
Tyr Leu Thr Ala Glu Val Leu Glu Leu Ala Gly Asn Ala Ser Lys
 80          85          90
Asp Leu Lys Val Lys Arg Ile Thr Pro Arg His Leu Gln Leu Ala
 95          100         105
Ile Arg Gly Asp Glu Leu Asp Ser Leu Ile Lys Ala Thr Ile
 110         115         120
Ala Gly Gly Gly Val Ile Pro His Ile His Lys Ser Leu Ile Gly
 125         130         135
Lys Lys Gly Gln Gln Lys Thr Ala
 140

```

&lt;210&gt; 520

<211> 197  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:399421.1.orf3:2000SEP08

<400> 520  
 Ser Leu Ala Leu Leu Glu Arg Val Cys Ser Phe Pro Leu Gly Arg  
   1                  5                  10                  15  
 Ile His Arg Leu Leu Arg Lys Gly Asn Tyr Ala Glu Arg Ile Gly  
                   20                  25                  30  
 Ala Gly Ala Pro Val Tyr Leu Ala Ala Val Leu Glu Tyr Leu Thr  
                   35                  40                  45  
 Ala His Thr Pro Cys Arg Lys Gln Ala Met Arg Leu Ala Ile Thr  
                   50                  55                  60  
 Lys Lys Leu Arg Ile Ile Ala Pro Pro Pro Ala Ala Gly Asp Pro  
                   65                  70                  75  
 Ala Met Met Arg Asn Ser His Lys Leu Leu Gly Gly Val Thr Thr  
                   80                  85                  90  
 Cys Pro Arg Ala Glu Ser Cys Leu Thr Val Gln Ala Gly Ala Ala  
                   95                  100                 105  
 Ala His Arg Arg Leu Arg Val Thr Thr Ile Lys Pro Gln Ala Ser  
                  110                 115                 120  
 Tyr Leu Arg Leu Ser Ile Asp Lys Lys Thr Ser Val Ser Glu Lys  
                  125                 130                 135  
 Thr Lys Arg Leu Phe Ser Glu Pro Pro Thr Val Leu Arg Glu Lys  
                  140                 145                 150  
 Ala Val Arg Tyr Gln Thr Leu Leu Leu Leu Thr Ser His His His  
                  155                 160                 165  
 Thr Arg Ser Cys Pro Lys Cys Pro Phe Asn Ala His Phe Gly Asn  
                  170                 175                 180  
 Val Arg Gly Leu Asn Phe Gln Gly Ala Phe Pro Ser Ala Ile Ser  
                  185                 190                 195  
 Leu Arg

<210> 521  
 <211> 100  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:816655.2.orf2:2000SEP08

<400> 521  
 Ala His Val Val Asn Ser Gln Asn Thr Ser Leu Leu Phe Tyr Gly  
   1                  5                  10                 15  
 Glu Gly Ala Tyr Val Thr Ser Arg Met Ser Pro Arg Ala Gly Leu  
                   20                  25                  30  
 Met Trp Arg Lys His Leu Ser Leu Leu Val Leu Arg Asp Phe Pro  
                   35                  40                  45  
 Leu Ala Ser Gln Glu Gly Ile Pro Val Asp Phe Asp Thr Gln  
                   50                  55                  60  
 Trp Pro Pro Leu Ala Gln Lys Ser Leu Trp Tyr Arg Lys Thr Asn  
                   65                  70                  75  
 Leu Leu Phe Thr Val Leu Phe Ser Leu Ser Ile Phe Gln His Arg  
                   80                  85                  90  
 Leu Asn Ser Leu Lys Pro Gln Thr Ser Gly  
                   95                 100

<210> 522  
 <211> 153  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:414732.1.orf1:2000SEP08

<400> 522  
 Trp Tyr Arg Arg Leu Leu Arg Glu Ser Gly Ser Thr Met Asp Ile  
     1                  5                  10                  15  
 Pro Val Pro Ser Ser Phe Asn Asp Val Gly Gln Asp Trp Arg Leu  
                   20                  25                  30  
 Arg His Phe Val Asp Gln Met Trp Tyr Glu Arg Glu Val Thr Phe  
                   35                  40                  45  
 Leu Glu Gln Trp Thr Gln Asp Leu His Thr Arg Val Val Leu Arg  
                   50                  55                  60  
 Ile Val Ser Ala His Ser Tyr Ala Ile Val Trp Val Asn Gly Val  
                   65                  70                  75  
 Asp Ala Leu Glu His Glu Gly Ser Thr Ser Pro Leu Thr Pro Thr  
                   80                  85                  90  
 Ser Val Ala Cys Ser Arg Trp Gly Pro Cys Pro Pro Ala Ser Ala  
                   95                  100                 105  
 Ser Leu Ser Pro Ser Ala Thr Cys Ser Ser Pro Pro Pro Cys His  
                  110                 115                 120  
 Gln Gly Ala Ser Ser Thr Trp Pro Thr Pro Pro Arg Gly Tyr His  
                  125                 130                 135  
 Pro Ala Ser Thr Ala Asp Thr His Leu Pro Val Pro Pro Arg Gly  
                  140                 145                 150  
 Ala Leu His

<210> 523  
 <211> 86  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:1140250.1.orf3:2000SEP08

<400> 523  
 Ile Thr Leu Met Trp Pro Gln Pro His Leu Pro Pro His Pro Met  
     1                  5                  10                  15  
 Met Ser Glu Lys Thr Arg Gln Asn Lys Leu Ala Glu Ala Lys Lys  
                   20                  25                  30  
 Lys Phe Thr Asp Tyr Arg Gln Trp Asn Ile Ala Gly Val Gly Thr  
                   35                  40                  45  
 Gly Ala Thr Asp Thr Lys Lys Lys Lys Ile Asn His Gly Thr Asn  
                   50                  55                  60  
 Pro Glu Thr Thr Thr Ser Gly Gly Cys His Ser Pro Glu Asp Thr  
                   65                  70                  75  
 Gln Gln Asn Arg Ala Gln Leu Lys Glu Val Thr  
                   80                  85

<210> 524  
 <211> 101  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature

<223> Incyte ID No: LG:174022.1.orf1:2000SEP08.

<400> 524

Gln	Pro	Leu	Thr	Ser	Val	Arg	Ser	Lys	His	Ile	Cys	Gly	Arg	Ala	
1				5					10					15	
Lys	Gly	Leu	Gly	Pro	Ser	Ser	Ser	Gln	Leu	Leu	Pro	Cys	His	Trp	
				20					25					30	
Val	Ala	Thr	Pro	Gly	Arg	Thr	Ile	Ile	Pro	Pro	Asp	Leu	Ser	Val	
				35					40					45	
Tyr	Pro	Gln	Gly	Leu	Thr	Phe	Pro	Ile	Gly	Pro	Ala	Gly	His	Tyr	
				50					55					60	
Ser	Pro	Gly	Gly	Ser	Arg	Arg	Leu	Ala	Thr	Glu	Val	Val	Lys	Ala	
				65					70					75	
Lys	Ala	Gly	Ala	Arg	Leu	Cys	His	Ser	Val	His	Gly	Leu	Cys	Trp	
				80					85					90	
Ser	Pro	Leu	Cys	Leu	Leu	Pro	Gly	Trp	Thr	Pro					
				95					100						

<210> 525

<211> 103

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:002811.1.orf1:2000SEP08

<400> 525

Val	Cys	Thr	Thr	Tyr	Leu	Ser	Phe	Cys	Pro	Gln	Glu	Pro	Gly	Leu	
1				5					10					15	
Pro	Val	Tyr	His	Pro	Ser	Val	Asp	Thr	Leu	Val	Ser	Asp	Asp	Ser	
				20					25					30	
Ala	Leu	Phe	Ala	Asn	Pro	Val	His	Tyr	Asp	Phe	Phe	Pro	Ser	Arg	
				35					40					45	
Ile	Asp	Leu	Thr	Gly	Leu	Lys	Phe	Tyr	Phe	His	Ile	Trp	Leu	Cys	
				50					55					60	
Ser	Ser	Asn	Ile	Leu	Pro	His	Trp	Thr	Cys	Gly	Ile	Val	Leu	Arg	
				65					70					75	
Phe	Arg	Glu	Gln	Gly	Glu	Asn	Arg	Ala	Gly	Pro	Pro	Leu	Pro	Leu	
				80					85					90	
Gly	Ala	Gln	Pro	Pro	Asn	Arg	Pro	Ser	His	Ser	Leu	Ser			
				95					100						

<210> 526

<211> 153

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: LI:414732.2.orf1:2000SEP08

<400> 526

Trp	Tyr	Arg	Arg	Leu	Leu	Arg	Glu	Ser	Gly	Ser	Thr	Met	Asp	Ile	
1				5					10					15	
Pro	Val	Pro	Ser	Ser	Phe	Asn	Asp	Val	Gly	Gln	Asp	Trp	Arg	Leu	
				20					25					30	
Arg	His	Phe	Val	Asp	Gln	Met	Trp	Tyr	Glu	Arg	Glu	Val	Thr	Phe	
				35					40					45	
Leu	Glu	Gln	Trp	Thr	Gln	Asp	Leu	His	Thr	Arg	Val	Val	Leu	Arg	
				50					55					60	
Ile	Val	Ser	Ala	His	Ser	Tyr	Ala	Ile	Val	Trp	Val	Asn	Gly	Val	
				65					70					75	

```

Asp Ala Leu Glu His Glu Gly Ser Thr Ser Pro Leu Thr Pro Thr
      80      85      90
Ser Val Ala Cys Ser Arg Trp Gly Pro Cys Pro Pro Ala Ser Ala
      95     100     105
Ser Leu Ser Pro Ser Ala Thr Cys Ser Ser Pro Pro Pro Cys His
      110     115     120
Gln Gly Ala Ser Ser Thr Trp Pro Thr Pro Pro Arg Gly Tyr His
      125     130     135
Pro Ala Ser Thr Ala Asp Thr His Leu Pro Val Pro Pro Arg Gly
      140     145     150
Ala Leu His

```

```

<210> 527
<211> 119
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:1019920.1.orf1:2000SEP08

```

```

<400> 527
Val Leu Arg Ser Gly Arg Asn Met Ala Ser Ala Thr Arg Val Ile
  1      5      10      15
Gln Lys Leu Arg Asn Trp Ala Ser Gly Gln Asp Leu Gln Ala Lys
      20      25      30
Leu Gln Leu Arg Tyr Gln Glu Ile Ala Lys Arg Thr Gln Pro Pro
      35      40      45
Pro Lys Leu Pro Val Gly Pro Ser His Lys Leu Ser Asn Asn Tyr
      50      55      60
Tyr Cys Thr Arg Asp Gly Arg Arg Glu Val Val Pro Pro Ser Ile
      65      70      75
Ile Met Ser Ser Gln Lys Ala Leu Val Ser Gly Lys Thr Ala Glu
      80      85      90
Ser Ser Ala Val Ala Ala Thr Lys Arg Ala Val Thr Pro Ala Pro
      95     100     105
Pro Met Lys Arg Trp Glu Leu Ser Arg Asp Gln Pro Tyr Leu
      110     115

```

```

<210> 528
<211> 62
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: LI:1038336.1.orf2:2000SEP08

```

```

<400> 528
Val Arg Leu Lys Lys Lys Phe Ser Gly Asn Phe Arg Met Leu Tyr
  1      5      10      15
Val Leu Val Ala Ile Ala Ala Ala Phe Leu Lys Asn Ala Trp Pro
      20      25      30
Lys Glu Arg Val Leu Val Leu Phe Leu Thr Ile Gly Gly Leu Ala
      35      40      45
Ile Asn Leu Thr Pro Val Cys Pro Tyr Thr Met Tyr Ser Thr Arg
      50      55      60
Ile Asn

```

```

<210> 529
<211> 381

```

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1177772.11.orf1:2000SEP08

&lt;400&gt; 529

Gly	Ser	Pro	His	Thr	Pro	Thr	Ser	Pro	Asn	Asp	Tyr	Ser	Leu	Tyr
1				5					10					15
Pro	Ser	Pro	Gln	Ser	Ser	Ser	Asn	Ser	Phe	Ser	Leu	His	Ala	Pro
			20						25					30
Gln	Ser	Gln	Tyr	Gln	Glu	Leu	Ala	Val	Ala	Leu	Asp	Ser	Ser	Ser
			35						40					45
Ala	Ile	Ile	Ser	Gln	Leu	Thr	Glu	Asn	Ile	Asn	Ser	Leu	Val	Arg
			50						55					60
Thr	Ser	Lys	Glu	Glu	Lys	Lys	His	Glu	Ile	His	Leu	Val	Gln	Lys
			65						70					75
Leu	Gly	Arg	Ser	Leu	Phe	Lys	Leu	Lys	Asn	Gln	Thr	Ala	Glu	Pro
			80						85					90
Leu	Ala	Pro	Gln	Pro	Pro	Ala	Gly	Pro	Ser	Lys	Val	Glu	Gln	Leu
			95						100					105
Gln	Asp	Glu	Thr	Asn	His	Leu	Arg	Lys	Glu	Leu	Glu	Ser	Val	Gly
			110						115					120
Arg	Gln	Leu	Gln	Ala	Glu	Val	Glu	Asn	Asn	Gln	Met	Leu	Ser	Leu
			125						130					135
Leu	Asn	Arg	Arg	Gln	Glu	Glu	Arg	Leu	Arg	Glu	Gln	Glu	Glu	Arg
			140						145					150
Leu	Arg	Glu	Gln	Glu	Glu	Arg	Leu	Cys	Glu	Gln	Glu	Glu	Arg	Leu
			155						160					165
Cys	Glu	Gln	Glu	Glu	Arg	Leu	Arg	Glu	Gln	Glu	Glu	Arg	Leu	Cys
			170						175					180
Glu	Gln	Glu	Lys	Leu	Pro	Gly	Gln	Glu	Arg	Leu	Leu	Glu	Glu	Val
			185						190					195
Glu	Lys	Leu	Leu	Glu	Gln	Glu	Arg	Arg	Gln	Glu	Glu	Gln	Glu	Arg
			200						205					210
Leu	Leu	Glu	Arg	Glu	Arg	Leu	Leu	Asp	Glu	Val	Glu	Glu	Leu	Leu
			215						220					225
Glu	Gln	Glu	Arg	Leu	Arg	Gln	Gln	Asp	Glu	Arg	Leu	Trp	Gln	Gln
			230						235					240
Glu	Thr	Leu	Arg	Glu	Leu	Glu	Arg	Leu	Arg	Glu	Leu	Glu	Arg	Leu
			245						250					255
Arg	Glu	Leu	Glu	Arg	Met	Leu	Glu	Leu	Gly	Trp	Glu	Ala	Leu	Tyr
			260						265					270
Glu	Gln	Arg	Ala	Glu	Pro	Arg	Ser	Gly	Phe	Glu	Glu	Leu	Asn	Asn
			275						280					285
Glu	Asn	Lys	Ser	Thr	Leu	Gln	Leu	Glu	Gln	Gln	Val	Lys	Glu	Leu
			290						295					300
Glu	Lys	Ser	Gly	Gly	Ala	Glu	Glu	Pro	Arg	Gly	Ser	Glu	Ser	Ala
			305						310					315
Ala	Ala	Ala	Arg	Pro	Val	Pro	Gly	Ala	Pro	Phe	Pro	Gln	Gly	Ala
			320						325					330
Trp	Met	Cys	Gly	Gln	Ala	Gly	Trp	Thr	Pro	Gln	Glu	His	Pro	Gly
			335						340					345
Leu	Ser	Gly	Glu	Ala	Val	Gly	Thr	Gly	Glu	Ala	Ala	Gly	Gly	Ala
			350						355					360
Gly	Glu	Ala	Ala	Cys	His	Ser	Phe	Arg	Ala	Ala	Glu	Asn	Arg	Glu
			365						370					375
Leu	Asn	Ile	Thr	Ile	Ile									
														380

&lt;210&gt; 530

&lt;211&gt; 173

<212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:205642.2.orf3:2000SEP08

<400> 530  
 Arg Arg Lys Pro Gln Thr Gln Pro His Leu Phe Ser Arg Pro Leu  
 1 5 10 15  
 Pro Pro Pro Arg Ser Val Gly Arg Ser Ala Met Thr Pro Gly Ser  
 20 25 30  
 Ser Ala Ala Gly Ser Gly Val Val Val Pro Arg Asn Phe Arg Leu  
 35 40 45  
 Leu Glu Glu Leu Glu Arg Gly Glu Lys Gly Ile Gly Asp Gly Thr  
 50 55 60  
 Val Ser Tyr Gly Met Asp Asp Ala Asp Asp Ile Tyr Met Arg Ser  
 65 70 75  
 Trp Thr Gly Thr Ile Ile Gly Pro His Asn Thr Val His Glu Gly  
 80 85 90  
 Arg Ile Tyr Gln Leu Lys Leu Phe Cys Asp Lys Asp Tyr Pro Glu  
 95 100 105  
 Lys Pro Pro Ser Val Arg Phe His Ser Arg Ile Asn Leu Thr Cys  
 110 115 120  
 Val Asn His Glu Thr Gly Val Val Asp Pro Lys Lys Phe Ser Val  
 125 130 135  
 Leu Gly Asn Trp Gln Arg Asp Tyr Ser Met Glu Tyr Ile Leu Thr  
 140 145 150  
 His Leu Lys Lys Glu Met Thr Ser Pro Gln Asn Arg Lys Leu Val  
 155 160 165  
 Gln Pro Pro Glu Gly Thr Phe Phe  
 170

<210> 531  
 <211> 124  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:449685.1.orf3:2000SEP08

<400> 531  
 Ser Ala Arg Glu Pro Phe Glu Pro Pro Thr Ser Pro Arg Leu Val  
 1 5 10 15  
 Pro His Leu Leu Arg Leu Leu Ala Arg Ala Ala Ala Arg Arg Leu  
 20 25 30  
 Pro Leu Val Ser Ala Pro Arg Ala Val Val His Pro Cys Ala Ala  
 35 40 45  
 Lys Met Thr Ala Gly Tyr Ile Val Gly Ser Leu Val Gly Ser Phe  
 50 55 60  
 Ala Ile Ala Tyr Leu Cys Asp Thr Phe Ile Ser Asp Lys Lys Ala  
 65 70 75  
 Phe Gly Gly Ser Thr Pro Lys Thr Val Ser Glu Lys Glu Trp Trp  
 80 85 90  
 Gln Ala Thr Asp Thr Lys Phe Gln Ala Trp Pro Arg Thr Ala Gly  
 95 100 105  
 Pro Pro Val Val Met Asn Pro Ile Ser Arg Gln Asn Phe Ile Val  
 110 115 120  
 Lys Ser Thr Glu

<210> 532



<211> 135  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:453922.1.orf1:2000SEP08

<400> 532  
 Glu Gln Gln Ile Lys Ser Val Ile Ser Ser Val Pro Thr Lys Pro  
 1 5 10 15  
 Gln Gln Pro Ser Phe Met Ser Asp Leu Asp Val Gln Leu Pro Ser  
 20 25 30  
 Ala Phe Asp Pro Phe Ala Glu Ala Asn Ala Glu Asp Ser Gly Ala  
 35 40 45  
 Gly Pro Gly Thr Lys Asp Tyr Val His Val Arg Ile Gln Gln Arg  
 50 55 60  
 Asn Gly Arg Lys Ser Leu Thr Thr Val Gln Gly Leu Lys Lys Glu  
 65 70 75  
 Phe Ser Tyr Asn Lys Ile Leu Lys Asp Leu Lys Lys Glu Phe Cys  
 80 85 90  
 Cys Asn Gly Thr Val Val Gln Asp Pro Glu Leu Gly Gln Val Ile  
 95 100 105  
 Gln Leu Gln Gly Asp Gln Arg Lys Asn Val Ala Thr Phe Leu Val  
 110 115 120  
 Gln Ala Gly Ile Ala Lys Lys Glu Asn Ile Lys Ile His Gly Phe  
 125 130 135

<210> 533  
 <211> 121  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LG:476342.3.orf2:2000SEP08

<400> 533  
 Ile Asp Arg Ser Pro Asp Leu Ile Asp Asn Ser Arg Ala Asp Gly  
 1 5 10 15  
 Arg Thr Arg Val Ser Gly Pro Val Arg Asn Ser Thr Asp Pro Ile  
 20 25 30  
 His Ala Gly Arg Arg Trp Ser Ser Ser Ser Pro Gly Arg Pro Cys  
 35 40 45  
 Ala Arg Ser Ser Trp Arg Cys Ser Cys Ser Arg Thr Thr Thr Ala  
 50 55 60  
 Gly Arg Arg Arg Arg Trp Trp Arg Arg Pro Gly Cys Ala Trp Ala  
 65 70 75  
 Arg Ala Ser Thr Thr Arg Ser Pro Ala Ser Pro Thr Ala Ser Ala  
 80 85 90  
 Ala Thr Ser Ala Ser Arg Arg Thr Ala Gly Gly Pro Pro Ala Thr  
 95 100 105  
 Ala Thr Ser Ala Thr Ala Gly Ala Arg Arg Arg Ala Lys Gln Ser  
 110 115 120  
 Ser

<210> 534  
 <211> 121  
 <212> PRT  
 <213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:476342.3.orf3:2000SEP08

&lt;400&gt; 534

Ile	Thr	Leu	Gly	Arg	Thr	Asp	Gly	Arg	Ala	Ser	Pro	Val	Pro	Ser
1				5					10					15
Val	Ile	Ala	Leu	Ile	Arg	Ser	Thr	Pro	Ala	Gly	Asp	Gly	Ala	His
				20					25					30
Gln	Val	Gln	Gly	Asp	Arg	Val	Arg	Ala	Pro	Pro	Gly	Ala	Ala	Pro
				35					40					45
Ala	Leu	Ala	Leu	Arg	Arg	Arg	Asp	Asp	Asp	Asp	Asp	Gly	Gly	Gly
				50					55					60
Gly	Pro	Gly	Val	His	Gly	Gln	Glu	Pro	Ala	Pro	Leu	Val	Pro	Leu
				65					70					75
His	Leu	Arg	Pro	Pro	Leu	Gln	Gln	Arg	Val	Arg	Gln	Gly	Gly	Arg
				80					85					90
Arg	Val	Asp	Arg	Arg	Leu	Leu	Pro	Pro	Pro	Leu	Leu	Gln	Val	Pro
				95					100					105
Glu	Gly	Val	Leu	Ser	Lys	Ala	Leu	Glu	Thr	Pro	Leu	Ala	Cys	Gln
				110					115					120

Asn

&lt;210&gt; 535

&lt;211&gt; 172

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:336801.1.orf1:2000SEP08

&lt;400&gt; 535

His	Arg	Lys	Met	Trp	Lys	Arg	Ser	Asp	His	Gln	Pro	Lys	Ile	Lys
1				5					10					15
Ala	Glu	Asp	Gly	Pro	Leu	Val	Gly	Gln	Phe	Glu	Val	Leu	Gly	Ser
				20					25					30
Val	Pro	Glu	Pro	Ala	Met	Pro	His	Pro	Leu	Glu	Leu	Ser	Glu	Phe
				35					40					45
Glu	Ser	Phe	Pro	Val	Phe	Gln	Asp	Ile	Arg	Leu	His	Ile	Arg	Glu
				50					55					60
Val	Gly	Ala	Gln	Leu	Val	Lys	Lys	Val	Asn	Ala	Val	Phe	Gln	Leu
				65					70					75
Asp	Ile	Thr	Lys	Asn	Gly	Lys	Thr	Ile	Leu	Arg	Trp	Thr	Ile	Asp
				80					85					90
Leu	Lys	Asn	Gly	Ser	Gly	Asp	Met	Tyr	Pro	Gly	Pro	Ala	Arg	Leu
				95					100					105
Pro	Ala	Asp	Thr	Val	Phe	Thr	Ile	Pro	Glu	Ser	Val	Phe	Met	Glu
				110					115					120
Leu	Val	Leu	Gly	Lys	Met	Asn	Pro	Gln	Lys	Ala	Phe	Leu	Ala	Gly
				125					130					135
Lys	Phe	Lys	Val	Ser	Gly	Lys	Val	Leu	Leu	Ser	Trp	Lys	Leu	Glu
				140					145					150
Arg	Val	Phe	Lys	Asp	Trp	Ala	Lys	Leu	Leu	Ser	Met	Gln	Lys	Tyr
				155					160					165
Gln	Gly	Arg	Thr	Ser	Leu	Arg								
				170										

&lt;210&gt; 536

&lt;211&gt; 103

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:449685.1.orf2:2000SEP08

&lt;400&gt; 536

Thr	Pro	His	Val	Ser	Ser	Pro	Arg	Ser	Pro	Phe	Ala	Ala	Ser	Pro
1				5					10					15
Cys	Pro	Arg	Arg	Arg	Ser	Pro	Phe	Ala	Pro	Arg	Leu	Arg	Thr	Ala
				20					25					30
Arg	Cys	Cys	Pro	Ser	Leu	Cys	Ser	Lys	Asn	Asp	Gly	Trp	Leu	His
				35					40					45
Cys	Trp	Leu	Thr	Gly	Arg	Ile	Leu	Cys	Leu	Leu	His	Thr	Cys	Val
				50					55					60
Thr	His	Leu	Ser	Leu	Thr	Arg	Arg	His	Leu	Glu	Val	Ala	Pro	Pro
				65					70					75
Arg	Leu	Phe	Leu	Arg	Arg	Ser	Gly	Gly	Lys	Pro	Gln	Thr	Pro	Ser
				80					85					90
Ser	Arg	Pro	Gly	Leu	Ala	Leu	Leu	Gly	His	Arg	Leu	Ser		
				95					100					

&lt;210&gt; 537

&lt;211&gt; 122

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:476342.1.orf1:2000SEP08

&lt;400&gt; 537

Gly	Arg	Thr	Arg	Val	Ser	Gly	Pro	Val	Arg	Asn	Asn	Thr	Asp	Pro
1				5					10					15
Ile	His	Ala	Gly	Arg	Arg	Trp	Ser	Ser	Ser	Ser	Pro	Gly	Arg	Pro
				20					25					30
Cys	Ala	Arg	Ser	Ser	Trp	Arg	Cys	Ser	Cys	Ser	His	Thr	Thr	Thr
				35					40					45
Ala	Gly	Arg	Arg	Arg	Arg	Trp	Trp	Arg	Arg	Pro	Gly	Cys	Ala	Trp
				50					55					60
Ala	Arg	Ala	Ser	Thr	Thr	Arg	Ser	Pro	Ala	Ser	Pro	Thr	Ala	Ser
				65					70					75
Ala	Ala	Thr	Ser	Ala	Ser	Arg	Arg	Thr	Ala	Gly	Gly	Pro	Pro	Ala
				80					85					90
Thr	Ala	Thr	Ser	Ala	Thr	Ala	Gly	Ala	Arg	Arg	Arg	Ala	Lys	Gln
				95					100					105
Ser	Ser	Ser	Asn	Thr	Leu	Gly	Leu	Pro	Glu	Leu	Asn	Ser	Ser	Ser
				110					115					120

Thr Lys

&lt;210&gt; 538

&lt;211&gt; 168

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1072804.1.orf2:2000SEP08

&lt;400&gt; 538

Ser	Tyr	Leu	Arg	Ser	Arg	Gly	Gln	Pro	Pro	Pro	Arg	Arg	Ser	His
1				5					10					15
Ala	Leu	Arg	Ala	Arg	Arg	Leu	Ser	Ser	Val	Ser	Ala	Ser	Leu	Pro
				20					25					30

```

Leu Pro Ser Arg Leu Thr His Met Ala Ser Ile Ala Gly Ser Ser
      35      40
Ala Leu Ser Phe Ala Arg Pro Val Lys Ala Ile Asn Thr Asn Ser
      50      55      60
Leu Ala Phe Ser Pro Ala Arg Lys Gly Asn Thr Phe Leu Arg Leu
      65      70      75
Gln Pro Met Pro Met Arg Ser Val Ser Cys Ala Ala Lys Lys Asp
      80      85      90
Thr Thr Asp Lys Val Cys Glu Ile Val Lys Lys Gln Leu Ala Leu
      95     100     105
Pro Asp His Thr Glu Val Cys Gly Glu Ser Lys Phe Ser Glu Leu
     110     115     120
Gly Ala Asp Ser Leu Asp Thr Val Glu Ile Val Met Ser Leu Glu
     125     130     135
Glu His Phe Asp Ile Ser Val Glu Glu Ser Ser Ala Gln Thr Ile
     140     145     150
Ala Thr Val Glu Asp Ala Ala Asp Leu Ile Asp Lys Leu Val Ala
     155     160     165
Gly Lys Ala

```

&lt;210&gt; 539

&lt;211&gt; 142

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:455450.1.orf1:2000SEP08

&lt;400&gt; 539

```

Gly Gly Gly Ile Leu Arg Arg Ser Gly Ser Ser Ser Ser Ser Ser
  1      5      10
Ser Ser Ser Ser Glu Asp Asp Gly Met Gly Gly Arg Arg Lys Lys
  20      25      30
Gly Leu Lys Glu Lys Ile Lys Glu Lys Met Pro Gly Gly His Arg
  35      40      45
Glu Gly Gln Gly Gln Ala Thr Ala Thr Gly Ala Tyr Gly Gly Thr
  50      55      60
Gly Tyr Val Ala Gly Pro Thr Thr Gly Gly Pro His Glu Lys Lys
  65      70      75
Gly Val Val Glu Lys Ile Lys Glu Lys Ile Pro Gly Gly His Lys
  80      85      90
Asp Tyr Asp Gln His Gln His Thr Thr Ala Ala Thr Gly Gly Gly
  95     100     105
Gly Gly Tyr Gly Gly Thr Thr Asp Thr Tyr Gly Thr Thr Thr
  110     115     120
Thr Glu Gly Thr His Glu Lys Lys Gly Phe Met Asp Lys Ile Lys
  125     130     135
Glu Lys Leu Pro Gly Gln His
  140

```

&lt;210&gt; 540

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1073699.1.orf1:2000SEP08

&lt;400&gt; 540

```

Ala Cys Leu Glu Glu Leu Ser Tyr Ser Leu Ser Arg Leu Leu Ala

```

1	5	10	15
Trp Ile Trp Thr	Leu Arg Pro Ala Pro Val	Leu Leu Val Val	Pro
	20	25	30
Ala Pro Ala Arg	Thr Asn Ala Asn Ala Arg	Ala Ala Asn Ala	Arg
	35	40	45
Thr Ala Arg Arg	Ala Ala Ala Pro Val	Ala Leu Gln Asp Val	Arg
	50	55	60
Ser Val Pro Arg	Thr Val Phe Ala Lys	Ala Lys Arg Gly Pro	Arg
	65	70	75
Pro Arg Asn Ala	Ala Ala Ala Ser Glu Asp	Ser His Thr Ala	Tyr
	80	85	90
Val Asn Ser Ala	Ala Cys Pro Trp Trp Gly	Val Ala Val Ala	Pro
	95	100	105
Leu Pro Gly Phe	Leu Leu Arg Gly Cys	Glu	
	110	115	

&lt;210&gt; 541

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1013729.1.orf3:2000SEP08

&lt;400&gt; 541

Ala Leu Pro Leu	Ala Met Ala Ser Leu Lys	Asp Leu Glu Gly Lys
1	5	10
		15
Trp Arg Leu Val	Glu Ser His Gly Phe Glu	Asp Tyr Met Lys Glu
	20	25
		30
Leu Gly Val Gly	Leu Ala Leu Arg Lys Met	Gly Ala Met Ala Lys
	35	40
		45
Pro Asp Cys Ile	Ile Thr Leu Asp Gly Asn	Asn Leu Thr Val Lys
	50	55
		60
Thr Glu Ser Thr	Val Lys Thr Thr Val Phe	Ser Cys Thr Leu Gly
	65	70
		75
Glu Lys Phe Asp	Glu Thr Thr Ala Asp Gly	Arg Lys Thr Glu Thr
	80	85
		90
Val Cys Thr Phe	Thr Asp Gly Ala Leu Val	Gln His Gln Lys Trp
	95	100
		105
Glu Gly Lys Glu	Ser Thr Ile Thr Arg Lys	Leu Lys Asp Gly Lys
	110	115
		120
Met Val Val Glu	Cys Val Met Asn Asn Cys	Pro Ser Cys Thr Ser
	125	130
		135
Gly Leu Met Thr	Lys Val Gln Leu Arg Thr	Gly
	140	145

&lt;210&gt; 542

&lt;211&gt; 142

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2050322.2.orf2:2000SEP08

&lt;400&gt; 542

Thr Gly Ala Ser	Trp Ala Pro Ser His Ser	Thr Trp His Trp Pro
1	5	10
		15
Ser Ala Lys Glu	Glu Arg Lys Ala Ile Leu	Ser Asn Gln Tyr Met
	20	25
		30
Gln Arg Leu Ser	Thr Met Arg Thr Pro Glu	Gln Ser Pro Ser Trp
	35	40
		45

```

Ala Pro Phe Ser Ser Pro Pro Ala Thr Ser Cys Leu Pro Cys Pro
      50      55      60
Ser Cys Leu Pro Ser Arg Val Pro Gln Trp Ser Gly Gln Pro Gly
      65      70      75
Cys Gly Leu Gly Ala Pro Arg Pro Thr Phe Ser Ser Val Arg Gln
      80      85      90
Ala Ser Thr Gln Val Pro Arg Thr Val Pro His Thr Gln Arg Val
      95     100     105
Ala Asn Ile Gly Thr Gln Thr Thr Gly Pro Ser Gly Val Gly Cys
     110     115     120
Cys Thr Pro Gly Arg Pro Leu Ser Ala Val Gln Met Phe Leu Ser
     125     130     135
Ser Thr Tyr Ala Pro Ile Gly
      140

```

&lt;210&gt; 543

&lt;211&gt; 175

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:891327.1.orf3:2000SEP08

&lt;400&gt; 543

```

Lys Lys Cys Ile Leu Leu His Leu His Ala Pro Gln Ala Thr Val
  1      5      10      15
Ser Gln Leu Lys Glu Gln Arg Arg Leu Leu Asp Leu Gln Lys Arg
      20      25      30
Lys Lys Pro Ser Glu Glu Thr Gly Thr Lys Arg Ser Lys Met
      35      40      45
Ser Lys Glu Gln Thr Arg Pro Ser Cys Ser Ala Gly Ala Ser Thr
      50      55      60
Ser Thr Ala Met Gly Arg Ser Pro Pro Pro Gln Thr Ser Ser Ser
      65      70      75
Ala Pro Pro Asn Thr Ser Ser Thr Glu Ser Leu Lys Pro Leu Ala
      80      85      90
Asn Arg His Ala Thr Ala Ser Lys Asn Ile Phe Arg Glu Asp Pro
      95     100     105
Ile Ile Ala Met Val Leu Asn Ala Thr Lys Val Phe Lys Tyr Glu
     110     115     120
Ser Ser Glu Asn Glu Gln Arg Arg Met Phe His Ala Thr Val Ala
     125     130     135
Thr Gln Thr Gln Phe Phe His Val Lys Val Leu Asn Ile Asn Leu
     140     145     150
Lys Arg Lys Phe Ile Lys Lys Arg Ile Ile Ile Ser Asn Tyr
     155     160     165
Ser Lys Arg Asn Ser Leu Tyr Arg Gly Glu
     170     175

```

&lt;210&gt; 544

&lt;211&gt; 96

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2053076.1.orf3:2000SEP08

&lt;400&gt; 544

```

Thr Leu Ala Cys Ser Glu Ala Ala Glu Ser Gly Lys Lys Leu Ser
  1      5      10      15
Leu Phe Cys His Leu Leu Lys Pro Asp Tyr Cys Val Lys Gln Ala

```

	20		25		30									
Met	Lys	Ala	Ile	Leu	Thr	Asp	Gln	Pro	Met	Ile	Cys	Thr	Pro	Arg
	35		40		45									
Leu	Met	Tyr	Ile	Val	Thr	Phe	Met	Lys	Ser	Ile	Leu	Pro	Phe	Glu
	50		55		60									
Ala	Val	Val	Cys	Met	Tyr	Arg	Phe	Leu	Gly	Ala	Asp	Lys	Cys	Met
	65		70		75									
Tyr	Pro	Phe	Ile	Ala	His	Ile	Glu	Ser	Lys	Pro	Gln	Thr	Ile	Met
	80		85		90									
Lys	Gln	Ile	Asn	Gly	Ile									
	95													

&lt;210&gt; 545

&lt;211&gt; 139

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:220085.1.orf3:2000SEP08

&lt;400&gt; 545

Gly	Gly	Pro	Gly	Ser	Gly	Ala	Pro	Asn	Glu	Pro	Met	Val	Arg	Met
1				5					10					15
Thr	Val	Arg	Ser	Lys	Asp	Ile	Glu	Phe	Pro	Val	Asp	Thr	Tyr	Pro
				20					25					30
Glu	His	Ser	Val	Val	Thr	Thr	Pro	Val	Thr	Pro	Leu	Ser	Lys	Lys
				35					40					45
Thr	Ile	Asp	Ile	Ile	Lys	Ala	Thr	Gly	Val	Leu	Lys	Lys	Gln	Ala
				50					55					60
Phe	Tyr	Leu	Pro	Arg	Thr	Cys	Thr	Val	Gly	Arg	His	Glu	Val	Ile
				65					70					75
His	Gln	Phe	Leu	Tyr	Met	Pro	Asp	Cys	Pro	Leu	Pro	Leu	Leu	Gly
				80					85					90
Arg	Asp	Leu	Leu	Ser	Lys	Leu	Arg	Ala	Thr	Ile	Ser	Phe	Thr	Lys
				95					100					105
His	Ser	Ser	Leu	Pro	Leu	Lys	Leu	Pro	Gly	Met	Gly	Val	Ile	Met
				110					115					120
Ala	Leu	Thr	Val	Pro	Trp	Glu	Glu	Glu	Trp	Thr	Phe	Leu	Asn	Ser
				125					130					135
Gln	Ala	Lys	Arg											

&lt;210&gt; 546

&lt;211&gt; 219

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:406709.1.orf3:2000SEP08

&lt;400&gt; 546

Gln	Leu	Leu	Glu	Ala	Pro	Pro	Gly	Pro	Thr	Ser	Thr	Arg	His	Leu
1				5					10					15
Glu	Pro	Leu	Ala	Pro	Met	Asp	Pro	Leu	Arg	Arg	Ser	Pro	Ser	Pro
				20					25					30
Cys	Leu	Ser	Ser	Gln	Pro	Ser	Ser	Pro	Ser	Thr	Pro	Pro	Cys	Glu
				35					40					45
Met	Leu	Gly	Pro	Val	Gly	Ile	Glu	Ala	Val	Leu	Asp	Gln	Leu	Lys
				50					55					60
Ile	Lys	Ala	Met	Lys	Met	Gly	Phe	Glu	Phe	Asn	Ile	Met	Val	Val
				65					70					75

Gly	Gln	Ser	Gly	Leu	Gly	Lys	Ser	Thr	Met	Val	Asn	Thr	Leu	Phe
				80					85					90
Lys	Ser	Lys	Val	Trp	Lys	Ser	Asn	Pro	Pro	Gly	Leu	Gly	Val	Pro
				95					100					105
Thr	Pro	Gln	Thr	Leu	Gln	Leu	His	Ser	Leu	Thr	His	Val	Ile	Glu
				110					115					120
Glu	Lys	Gly	Val	Lys	Leu	Lys	Leu	Thr	Val	Thr	Asp	Thr	Pro	Gly
				125					130					135
Phe	Gly	Asp	Gln	Ile	Asn	Asn	Asp	Asn	Cys	Trp	Asp	Pro	Ile	Leu
				140					145					150
Gly	Tyr	Ile	Asn	Glu	Gln	Tyr	Glu	Gln	Tyr	Leu	Gln	Glu	Glu	Ile
				155					160					165
Leu	Ile	Thr	Arg	Gln	Arg	His	Ile	Pro	Asp	Thr	Arg	Val	His	Cys
				170					175					180
Cys	Val	Tyr	Phe	Val	Pro	Pro	Thr	Gly	His	Cys	Leu	Arg	Pro	Leu
				185					190					195
Asp	Ile	Glu	Phe	Leu	Gln	Arg	Leu	Cys	Pro	Asp	Cys	Glu	Cys	Gly
				200					205					210
Ala	Arg	Asp	Cys	Gln	Gly	Gln	Thr	Ala						
				215										

&lt;210&gt; 547

&lt;211&gt; 161

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LG:347863.9.orf3:2000SEP08

&lt;400&gt; 547

Ser	Ala	Gly	Ile	Trp	Met	Phe	Met	Cys	Cys	Val	Cys	Val	Cys	Ala
1				5					10					15
His	Thr	Leu	Cys	Leu	Phe	Val	Cys	Val	His	Val	Tyr	Thr	Leu	Cys
				20					25					30
Ala	His	Val	Ser	Gly	Val	Cys	Val	Pro	Cys	Leu	Ser	Met	Tyr	Gly
				35					40					45
Gly	Ala	His	Val	Ala	Cys	Glu	Phe	Leu	Phe	Cys	Trp	Ala	Gly	Val
				50					55					60
Trp	Arg	Trp	Trp	Gly	Arg	Leu	Ser	Met	Lys	Gly	His	Leu	Cys	Gln
				65					70					75
Leu	Glu	Glu	Gly	Gly	Arg	Cys	Cys	Leu	Leu	Pro	Gln	Asp	Leu	Val
				80					85					90
Gly	Leu	Trp	Glu	Ala	Gly	Leu	His	Pro	Pro	Ala	Phe	Pro	Val	Leu
				95					100					105
Gly	Pro	Gln	Phe	Lys	Arg	Arg	Val	Gln	Glu	Ser	Thr	Gln	Val	Leu
				110					115					120
Arg	Glu	Leu	Glu	Thr	Ser	Leu	Arg	Thr	Asn	His	Ile	Gly	Trp	Val
				125					130					135
Gln	Glu	Phe	Leu	Asn	Glu	Glu	Asn	Arg	Gly	Leu	Asp	Val	Leu	Leu
				140					145					150
Glu	Tyr	Leu	Ala	Phe	Ala	Gln	Cys	Ser	Val	Thr				
				155					160					

&lt;210&gt; 548

&lt;211&gt; 177

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:1073027.1.orf1:2000SEP08



&lt;400&gt; 548

```

Ile Glu Met Ala Ser Lys Arg Ala Leu Val Ile Leu Ala Lys Gly
 1           5           10           15
Ala Glu Glu Met Glu Thr Val Ile Pro Val Asp Ile Met Arg Arg
           20           25           30
Ala Gly Ile Lys Val Thr Val Ala Gly Leu Ala Gly Lys Asp Pro
           35           40           45
Val Gln Cys Ser Arg Asp Val Val Ile Cys Pro Asp Thr Ser Leu
           50           55           60
Glu Glu Ala Lys Thr Gln Gly Pro Tyr Asp Val Val Val Leu Pro
           65           70           75
Gly Gly Asn Leu Gly Ala Gln Asn Leu Ser Glu Ser Ala Leu Val
           80           85           90
Lys Glu Ile Leu Lys Glu Gln Glu Asn Arg Lys Gly Leu Ile Ala
           95          100          105
Ala Ile Cys Ala Gly Pro Thr Ala Leu Leu Ala His Glu Val Gly
          110          115          120
Phe Gly Cys Lys Val Thr Ser His Pro Leu Ala Lys Asp Lys Met
          125          130          135
Met Asn Gly Ser His Tyr Ser Tyr Ser Glu Ser Arg Val Glu Lys
          140          145          150
Asp Gly Leu Ile Leu Thr Ser Arg Gly Pro Gly Thr Ser Phe Glu
          155          160          165
Phe Ala Leu Ala Ile Val Asp Pro Gly Pro Pro Gly
          170          175

```

&lt;210&gt; 549

&lt;211&gt; 241

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:347635.1.orf2:2000SEP08

&lt;400&gt; 549

```

Pro Arg Gly Lys His Val Ser Cys Val Ile Trp Met Ile Ala Cys
 1           5           10           15
Ile Ser Asn Pro Cys His Lys Gly Ala Leu Cys Asp Thr Asn Pro
           20           25           30
Leu Asn Gly Gln Tyr Ile Cys Thr Cys Pro Gln Gly Tyr Lys Gly
           35           40           45
Ala Asp Cys Thr Glu Asp Val Asp Glu Cys Ala Met Ala Asn Ser
           50           55           60
Asn Pro Cys Glu His Ala Gly Lys Cys Val Asn Thr Asp Gly Ala
           65           70           75
Phe His Cys Glu Cys Leu Lys Gly Tyr Ala Gly Pro Arg Cys Glu
           80           85           90
Met Asp Ile Asn Glu Cys His Ser Asp Pro Cys Gln Asn Asp Ala
           95          100          105
Thr Cys Leu Asp Lys Ile Gly Gly Phe Thr Cys Leu Cys Met Pro
          110          115          120
Gly Phe Lys Gly Val His Cys Glu Leu Glu Ile Asn Glu Cys Gln
          125          130          135
Ser Asn Pro Cys Val Asn Asn Gly Gln Cys Val Asp Lys Val Asn
          140          145          150
Arg Phe Gln Cys Leu Cys Pro Pro Gly Phe Thr Gly Pro Val Cys
          155          160          165
Gln Ile Asp Ile Asp Asp Cys Ser Ser Thr Pro Cys Leu Asn Gly
          170          175          180
Ala Lys Cys Ile Asp His Pro Asn Gly Tyr Glu Cys Gln Cys Ala
          185          190          195
Thr Gly Phe Thr Gly Val Leu Cys Glu Glu Asn Ile Asp Asn Cys

```

	200		205		210
Asp Pro Asp Pro	Cys His His Gly Gln	Cys Gln Asp Gly Ile Asp			
	215		220		225
Ser Tyr Thr Cys	Ile Cys Asn Pro Gly	Tyr Met Gly Ala Asn Leu			
	230		235		240
Gln					

<210> 550  
 <211> 108  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:013685.1.orf2:2000SEP08

<400> 550

Asp Val Val Ala Cys Gly Gly Ser Ser Pro Ile Gln Cys Met His	
1 5 10 15	
Pro Ala Pro Leu Phe Arg Leu Tyr Gly Pro Cys Gly Leu Arg Gln	
20 25 30	
Gly Pro Asp Thr Tyr Met Cys Val Glu Ile Arg Ser Leu Leu Ser	
35 40 45	
Leu Ser Cys His Lys Ser Gly Gly Glu Cys Pro Gly Pro Ser Val	
50 55 60	
Gly Ser Leu Ser Gly Val Cys Ser Leu His Pro Ser Trp Asn Leu	
65 70 75	
Pro Met Val Arg Arg Ser Arg Ser Ser Glu Pro Ser Ala Leu Val	
80 85 90	
Ser Pro Ile Gln Ser Ile Gly Arg Arg Gln Thr Leu Phe Pro Thr	
95 100 105	
Pro Pro Ser	

<210> 551  
 <211> 236  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: LI:406709.1.orf3:2000SEP08

<400> 551

Gln Leu Leu Glu Ala Pro Pro Gly Pro Thr Ser Thr Arg His Leu	
1 5 10 15	
Glu Pro Leu Ala Pro Met Asp Pro Leu Arg Arg Ser Pro Ser Pro	
20 25 30	
Cys Leu Ser Ser Gln Pro Ser Ser Pro Ser Thr Pro Pro Cys Glu	
35 40 45	
Met Leu Gly Pro Val Gly Ile Glu Ala Val Leu Asp Gln Leu Lys	
50 55 60	
Ile Lys Ala Met Lys Met Gly Phe Glu Phe Asn Ile Met Val Val	
65 70 75	
Gly Gln Ser Gly Leu Gly Lys Ser Thr Met Val Asn Thr Leu Phe	
80 85 90	
Lys Ser Lys Val Trp Lys Ser Asn Pro Pro Gly Leu Gly Val Pro	
95 100 105	
Thr Pro Gln Thr Leu Gln Ala Ala Phe Thr Asp Pro Cys His Arg	
110 115 120	
Gly Glu Gly Cys Glu Ala Glu Ala Asp Gly Asp Arg Thr Arg Pro	
125 130 135	

Ala	Ser	Gly	Thr	Arg	Ser	Thr	Met	Thr	Thr	Ala	Gly	Thr	Pro	Ser	
				140					145					150	
Trp	Ala	Thr	Ser	Thr	Ser	Asn	Thr	Ser	Ser	Thr	Cys	Arg	Arg	Arg	
				155					160					165	
Ser	Ser	Ser	Pro	Ala	Ser	Ala	Thr	Ser	Gln	Thr	Pro	Gly	Cys	Thr	
				170					175					180	
Ala	Ala	Cys	Thr	Leu	Tyr	His	Pro	Leu	Gly	Thr	Ala	Cys	Gly	Pro	
				185					190					195	
Trp	Thr	Leu	Ser	Ser	Cys	Ser	Gly	Cys	Val	Arg	Thr	Val	Asn	Val	
				200					205					210	
Val	Pro	Val	Ile	Ala	Arg	Ala	Arg	Gln	Pro	Asp	Pro	Trp	Arg	Ser	
				215					220					225	
Glu	Arg	Pro	Ser	Gly	Ala	Gly	Ser	Ser	Arg	Thr					
				230					235						

&lt;210&gt; 552

&lt;211&gt; 137

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:2052938.1.orf1:2000SEP08

&lt;400&gt; 552

Gly	Met	Val	Ser	Asn	Leu	Lys	Arg	Ile	Met	Ala	Gln	Lys	Trp	Met	
1				5					10					15	
Pro	Asp	Gln	Ile	Ser	Val	Ser	Glu	Phe	Ile	Ala	Glu	Thr	Thr	Glu	
				20					25					30	
Asp	Tyr	Asn	Ser	Pro	Thr	Thr	Ser	Ser	Phe	Thr	Thr	Arg	Leu	His	
				35					40					45	
Asn	Cys	Arg	Asn	Thr	Val	Thr	Leu	Leu	Glu	Glu	Ala	Leu	Glu	Gln	
				50					55					60	
Asp	Arg	Thr	Ala	Leu	Gln	Lys	Val	Lys	Lys	Ser	Val	Lys	Ala	Ile	
				65					70					75	
Tyr	Asn	Ser	Gly	Gln	Asp	His	Val	Gln	Asn	Glu	Glu	Asn	Tyr	Ala	
				80					85					90	
Gln	Val	Leu	Asp	Lys	Phe	Gly	Ser	Asn	Phe	Leu	Ser	Arg	Asp	Asn	
				95					100					105	
Pro	Ala	Pro	Trp	His	Arg	Val	Cys	Gln	Val	Phe	Tyr	Ser	Tyr	Lys	
				110					115					120	
Gly	Thr	Val	His	Thr	Ala	Glu	Lys	Ser	Ala	Pro	Gly	Phe	Glu	Pro	
				125					130					135	
Thr	Met														

&lt;210&gt; 553

&lt;211&gt; 150

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: LI:213208.1.orf3:2000SEP08

&lt;400&gt; 553

Val	Gly	Val	Cys	Ala	Glu	Arg	Ser	Gly	Pro	Arg	Ala	Gly	Gly	Arg	
1				5					10					15	
Pro	Arg	Val	Phe	Gly	Ser	Arg	Pro	Gln	Gly	Ala	Glu	Arg	Ser	Trp	
				20					25					30	
Asp	Arg	Arg	Pro	Pro	Leu	Leu	Pro	Gly	Met	Ser	Ala	Glu	Ala	Ser	
				35					40					45	
Gly	Pro	Ala	Ala	Ala	Ala	Ala	Pro	Ser	Leu	Glu	Ala	Pro	Lys	Pro	

	50		55		60
Ser Gly Leu Glu	Pro Gly Pro Ala Ala	Tyr Gly Leu Lys Pro	Leu		
	65		70		75
Thr Pro Asn Ser Lys	Tyr Val Lys Leu Asn Val Gly Gly Ser	Leu			
	80		85		90
His Tyr Thr Thr	Leu Arg Thr Leu Thr Gly Gln Asp Thr Met	Leu			
	95		100		105
Lys Ala Met Phe Ser Gly Arg Val Glu	Val Leu Thr Asp Ala Gly				
	110		115		120
Gly Trp Val Leu Ile Asp Arg Ser Gly	Arg His Phe Gly Thr Ile				
	125		130		135
Leu Asn Tyr Leu Arg Asp Gly Ser Val	Pro Leu Pro Glu Ser Thr				
	140		145		150

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number  
WO 02/020754 A3(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
5/10, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, A61K  
38/00

(21) International Application Number: PCT/US01/27127

(22) International Filing Date: 29 August 2001 (29.08.2001)

(25) Filing Language: English

(26) Publication Language: English

## (30) Priority Data:

60/229,751	5 September 2000 (05.09.2000)	US
60/229,749	5 September 2000 (05.09.2000)	US
60/229,750	5 September 2000 (05.09.2000)	US
60/229,747	5 September 2000 (05.09.2000)	US
60/229,748	5 September 2000 (05.09.2000)	US
60/230,583	5 September 2000 (05.09.2000)	US
60/230,517	6 September 2000 (06.09.2000)	US
60/230,599	6 September 2000 (06.09.2000)	US
60/230,514	6 September 2000 (06.09.2000)	US
60/230,598	6 September 2000 (06.09.2000)	US
60/230,988	6 September 2000 (06.09.2000)	US
60/230,518	6 September 2000 (06.09.2000)	US
60/230,515	6 September 2000 (06.09.2000)	US
60/230,597	6 September 2000 (06.09.2000)	US
60/230,505	6 September 2000 (06.09.2000)	US
60/230,519	6 September 2000 (06.09.2000)	US
60/230,595	6 September 2000 (06.09.2000)	US
60/230,865	6 September 2000 (06.09.2000)	US
60/230,610	6 September 2000 (06.09.2000)	US
60/231,167	7 September 2000 (07.09.2000)	US
60/231,163	7 September 2000 (07.09.2000)	US
60/230,951	7 September 2000 (07.09.2000)	US

(71) Applicant (for all designated States except US): INCYTE  
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94304 (US).

## (72) Inventors; and

(75) Inventors/Applicants (for US only): STUART, Jack-  
son [US/US]; 4913 Deep Creek Road, Fremont, CA  
94555 (US). LINCOLN, Stephen, E. [US/US]; 725  
Sapphire Street, Redwood City, CA 94061 (US). ALTUS,  
Christina, M. [US/US]; 625 Virginia Avenue, Camp-  
bell, CA 95008 (US). DUFOUR, Gerard, E. [US/US];

5327 Greenridge Road, Castro Valley, CA 94552 (US).  
CHALUP, Michael, S. [US/US]; 183 Acalanes Drive,  
Apt. 6., Sunnyvale, CA 94086 (US). JACKSON, Jen-  
nifer, L. [US/US]; 1826 Rina Court, Santa Cruz, CA  
95062 (US). JONES, Anissa, Lee [US/US]; 445 South  
15th Street, San Jose, CA 95112 (US). YU, Jimmy, Y.  
[US/US]; 37330 Portico Terrace, Fremont, CA 94536  
(US). WRIGHT, Rachel, J. [NZ/US]; 333 Anna Avenue,  
Mountain View, CA 94043 (US). GIETZEN, Darryl  
[US/US]; 691 Los Huecos Drive, San Jose, CA 95123  
(US). LIU, Tommy, F. [US/US]; 201 Ottilia Street, Daly  
city, Ca 94014 (US). YAP, Pierre, E. [US/US]; 201  
Happy Hollow Court, Lafayette, CA 94549 (US). DAHL,  
Christopher, R. [US/US]; 41277 Roberts Avenue, #6,  
Fremont, CA 94538 (US). MOMIYAMA, Monika, G.  
[US/US]; 25689 Deer Trail Place, Hayward, CA 94541  
(US). BRADLEY, Diana, L. [US/US]; 3260 Crystal  
Heights Drive, Soquel, CA 95073 (US). ROHATGI,  
Sameer, D. [US/US]; 5 Rico Way #103, San Francisco,  
CA 94123 (US). HARRIS, Bernard [US/US]; 1014  
Lupine Drive, Sunnyvale, CA 94086 (US). ROSEBERRY  
LINCOLN, Ann, M. [US/US]; 10637 Rock Run Drive,  
Potomac, MD 20854 (US). GERSTIN, Edward, H.,  
Jr. [US/US]; 1408 38th Avenue, San Francisco, CA  
94122 (US). PERALTA, Careyna, H. [US/US]; 4585  
Lakeshore Drive, Santa Clara, CA 95054 (US). DAVID,  
Marie, H. [US/US]; 131 Mirada Drive, Daly City, CA  
94015 (US). PANZER, Scott, R. [US/US]; 965 East El  
Camino, #621, Sunnyvale, CA 94087 (US). FLORES,  
Vincent [US/US]; 35000 Begonia Street, Union City, CA  
94587 (US). DAFFO, Abel [US/US]; 1750 Stokes Street,  
#70, San Jose, CA 95126 (US). MARWAHA, Rakesh  
[US/US]; 16272 Saratoga Street, #4, San Leandro, CA  
94578 (US). CHEN, Alice, J. [US/US]; 4405 Norwalk  
Drive, #22, San Jose, CA 95129 (US). CHANG, Simon,  
C. [US/US]; 1901 Rock Street, #103, Mountain View, CA  
94043 (US). AU, Alan, P. [US/US]; 565 Ortega Avenue,  
#3, Mountain View, CA 94040 (US). INMAN, Rebekah,  
R. [US/US]; 210 Calderon Avenue, #23, Mountain View,  
CA 94041 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,

[Continued on next page]

(54) Title: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS

(57) Abstract: The present invention provides purified human polynucleotides for diagnostics and therapeutics (dithp). Also en-  
compassed are the polypeptides (DITHP) encoded by dithp. The invention also provides for the use of dithp, or complements,  
oligonucleotides, or fragments thereof in diagnostic assays. The invention further provides for vectors and host cells containing  
dithp for the expression of DITHP. The invention additionally provides for the use of isolated and purified DITHP to induce antibod-  
ies and to screen libraries of compounds and the use of anti-DITHP antibodies in diagnostic assays. Also provided are microarrays  
containing dithp and methods of use.

WO 02/020754 A3



CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,  
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,  
ZW.

CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,  
TG).

**Published:**

— with international search report

**(88) Date of publication of the international search report:**

25 September 2003

**(84) Designated States (regional):** ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/27127

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C12N5/10 C07K14/47 C07K16/18 C12Q1/68  
G01N33/50 A61K38/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

WPI Data, EPO-Internal, SEQUENCE SEARCH, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] EBI Hinxton, UK; 30 November 1999 (1999-11-30) NCI.CGAP: "National Cancer Institute, Cancer Genome Anatomy Project (CGAP)" Database accession no. AW194769 XP002218725 abstract	1,3,4, 10-13,31
X	NISHI S ET AL: "HUMAN HEXOKINASE SEQUENCES OF AMINO AND CARBOXYL-TERMINAL HALVES ARE HOMOLOGOUS" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 157, no. 3, 1988, pages 937-943, XP002218724 ISSN: 0006-291X the whole document	29,31
	--- -/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&amp;" document member of the same patent family

Date of the actual completion of the international search

29 October 2002

Date of mailing of the international search report

05.02.2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 851 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Oderwald, H

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 01/27127

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 45712 A (HUMAN GENOME SCIENCES INC ;FENG PING (US); NI JIAN (US); ROSEN CRA) 15 October 1998 (1998-10-15) page 2 -page 5; claims 1-23 ---	5-9, 17-20
A	WO 99 25825 A (BOUGUELERET LYDIE ;GENSET SA (FR); DUCLERT AYMERIC (FR); DUMAS MIL) 27 May 1999 (1999-05-27)  page 1 -page 9; claims 1-18 ---	21-28, 35, 37-40, 51,53, 54,57, 59,60
A	WO 98 48274 A (MOORE KEITH J ;SMITHKLINE BEECHAM PLC (GB); DUNNINGTON DAMIEN J (U) 29 October 1998 (1998-10-29) abstract; claims 1-14 ---	
P,X	WO 01 55301 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); BARASH STEVEN C (US) 2 August 2001 (2001-08-02) page 1 -page 4; claims 1-24; examples 1-66 see SEQ ID NO: 62 and 1292 page 790 -page 1005 ---	1-54,57, 60
E	WO 02 08399 A (INCYTE GENOMICS INC ;THORNTON MICHAEL (US)) 31 January 2002 (2002-01-31) see SEQ ID NO: 17 and 37 (pp. 160-162 and 190/191) page 1 -page 71; claims 1-44; examples 1-19 ---	1-54,57, 60
E	WO 01 90325 A (MEYERS RACHEL A ;WILLIAMSON MARK (US); MILLENNIUM PHARM INC (US)) 29 November 2001 (2001-11-29) the whole document -----	1-54,57, 60



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 01/27127

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
**see FURTHER INFORMATION sheet PCT/ISA/210**
2. ☒ Claims Nos.: **55, 56, 58, 59**  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
**see FURTHER INFORMATION sheet PCT/ISA/210**
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

**see additional sheet**

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
**1-60 (all partially)**

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.1

Although claims 38 and 40 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claim 53 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

## Continuation of Box I.2

Claims Nos.: 55, 56, 58, 59

Present claims 55, 56, 58 and 59 relate to a product/compound defined by reference to a desirable characteristic or property, namely an agonist or antagonist of a polypeptide.

The claims cover all products/compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such products/compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product/compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: claims 1-60 partially

An isolated polynucleotide with SEQ ID NO: 1 and a corresponding polypeptide with SEQ ID NO: 276, a composition for detection, a method for detecting a target polynucleotide, a transformed cell, a transgenic organism, a method for producing a polypeptide, methods of screening for a test compound, a microarray, a method for generating a transcript image, a method for assessing toxicity, an array comprising different nucleotide molecules, an antibody, a method of diagnosing, methods of detecting and purifying a polypeptide, methods for treating a disease comprising said polynucleotide and polypeptide.

Invention 2-275: claims 1-60 partially

same as invention 1 but comprising a polynucleotide and corresponding polypeptide in the order given in claim 1 (invention 2 is limited to SEQ ID NO: 2 and 277; and invention 275 is limited to SEQ ID NO: 275 and 553).

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/27127

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9845712	A	15-10-1998	AU 6952998 A	30-10-1998
			EP 0974058 A2	26-01-2000
			JP 2001521383 T	06-11-2001
			WO 9845712 A2	15-10-1998
			US 2002058307 A1	16-05-2002
-----				
WO 9925825	A	27-05-1999	AU 753099 B2	10-10-2002
			AU 1049199 A	07-06-1999
			CA 2302644 A1	27-05-1999
			EP 1029045 A2	23-08-2000
			WO 9925825 A2	27-05-1999
			JP 2001523453 T	27-11-2001
			AU 1503099 A	05-07-1999
			CA 2311572 A1	24-06-1999
			EP 1037977 A2	27-09-2000
			WO 9931236 A2	24-06-1999
			JP 2002508182 T	19-03-2002
			AU 2294499 A	23-08-1999
			CA 2316182 A1	12-08-1999
			EP 1053318 A2	22-11-2000
			WO 9940189 A2	12-08-1999
			JP 2002502605 T	29-01-2002
			US 6312922 B1	06-11-2001
-----				
WO 9848274	A	29-10-1998	EP 0977988 A1	09-02-2000
			JP 2002511744 T	16-04-2002
			WO 9848274 A1	29-10-1998
-----				
WO 0155301	A	02-08-2001	AU 2950801 A	07-08-2001
			AU 3095801 A	07-08-2001
			AU 3645901 A	07-08-2001
			AU 3646001 A	07-08-2001
			AU 3646101 A	07-08-2001
			AU 3646201 A	07-08-2001
			AU 3646301 A	07-08-2001
			AU 3646401 A	07-08-2001
			AU 3646501 A	07-08-2001
			AU 3646601 A	07-08-2001
			AU 3794301 A	07-08-2001
			AU 3794401 A	07-08-2001
			AU 3794701 A	07-08-2001
			AU 3794901 A	07-08-2001
			AU 3795001 A	07-08-2001
			AU 3795101 A	07-08-2001
			AU 3795201 A	07-08-2001
			AU 3795301 A	07-08-2001
			AU 3795401 A	07-08-2001
			AU 3795501 A	07-08-2001
			AU 3795701 A	07-08-2001
			AU 3795801 A	07-08-2001
			AU 3972601 A	07-08-2001
			AU 3972701 A	07-08-2001
			AU 3972801 A	07-08-2001
			AU 4140201 A	07-08-2001
			AU 4140301 A	07-08-2001
			AU 4140401 A	07-08-2001
			AU 4140501 A	07-08-2001
			AU 4140601 A	07-08-2001

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/27127

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0155301	A		AU 4140701 A	07-08-2001
			AU 4140801 A	07-08-2001
			AU 4140901 A	07-08-2001
			AU 4141001 A	07-08-2001
			AU 4141101 A	20-08-2001
			AU 4141201 A	07-08-2001
			AU 4141301 A	07-08-2001
			AU 4141401 A	07-08-2001
			AU 4141501 A	07-08-2001
			AU 4141601 A	07-08-2001
			AU 4141701 A	07-08-2001
			AU 4141801 A	07-08-2001
			AU 4141901 A	07-08-2001
			AU 4313401 A	07-08-2001
			AU 4313501 A	07-08-2001
			AU 4313601 A	07-08-2001
			AU 4313701 A	14-08-2001
			AU 4526201 A	07-08-2001
			AU 4719001 A	07-08-2001
			AU 4719101 A	07-08-2001
-----				
WO 0208399	A	31-01-2002	AU 7795901 A	05-02-2002
			WO 0208399 A2	31-01-2002
-----				
WO 0190325	A	29-11-2001	AU 6336001 A	03-12-2001
			WO 0190325 A2	29-11-2001
			US 2002009779 A1	24-01-2002
-----				